

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:58:25 : Search time 12.48 Seconds  
(without alignments)  
102.779 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_201\_257

Perfect score: 302

Sequence: 1 TSTSPTRSMAPGAVHLPQPV.....STSFLLPMGSPPPARGSTGD 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 152146

Minimum DB seq length: 0

Maximum DB seq length: 57

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	27.5	15	US-08-221-583-56	Sequence 56, Appl
2	83	27.5	15	PCT-US95-04018-56	Sequence 56, Appl
3	82	27.2	15	US-08-221-583-58	Sequence 58, Appl
4	82	27.2	15	PCT-US95-04018-58	Sequence 58, Appl
5	81	26.8	15	US-08-221-583-57	Sequence 57, Appl
6	81	26.8	15	US-08-221-583-62	Sequence 62, Appl
7	81	26.8	15	PCT-US95-04018-57	Sequence 57, Appl
8	81	26.8	15	PCT-US95-04018-62	Sequence 62, Appl
9	80	26.5	15	US-08-221-583-55	Sequence 55, Appl
10	80	26.5	15	US-08-221-583-59	Sequence 59, Appl
11	80	26.5	15	PCT-US95-04018-55	Sequence 55, Appl
12	80	26.5	15	PCT-US95-04018-59	Sequence 59, Appl
13	79	26.2	15	US-08-221-583-61	Sequence 61, Appl
14	79	26.2	15	PCT-US95-04018-61	Sequence 61, Appl
15	78	25.8	15	US-08-221-583-60	Sequence 60, Appl
16	78	25.8	15	PCT-US95-04018-60	Sequence 60, Appl
17	69	22.8	15	US-08-221-583-54	Sequence 54, Appl
18	69	22.8	15	PCT-US95-04018-54	Sequence 54, Appl
19	59	19.5	45	US-08-361-920-19	Sequence 19, Appl
20	59	19.5	45	US-08-479-939-19	Sequence 19, Appl
21	59	19.5	45	US-08-483-432-19	Sequence 19, Appl
22	58	19.2	54	US-08-471-780C-44	Sequence 44, Appl
23	58	19.2	54	US-08-467-282B-44	Sequence 44, Appl
24	58	19.2	54	US-08-471-282A-44	Sequence 44, Appl
25	58	19.2	54	US-08-466-710C-44	Sequence 44, Appl
26	58	19.2	54	US-08-468-739C-44	Sequence 44, Appl
27	55	18.2	15	US-08-221-583-53	Sequence 53, Appl

28 55 18.2 15 5 PCT-US95-04018-53 Sequence 53, Appl  
29 55 18.2 46 3 US-08-856-074A-39 Sequence 39, Appl  
30 55 18.2 48 6 5171885-7 Patent No. 5171885  
31 55 18.2 48 6 5518916-7 Patent No. 5518916  
32 53.5 17.7 37 3 US-08-814-052-37 Sequence 37, Appl  
33 53.5 17.7 37 3 US-08-812-829-29 Sequence 29, Appl  
34 52.5 17.4 21 5 PCT-US92-10432-1 Sequence 1, Appl  
35 52.5 17.4 22 1 US-08-443-542-43 Sequence 43, Appl  
36 52.5 17.4 22 3 US-08-765-469-43 Sequence 43, Appl  
37 50.5 16.7 53 6 5422248-4 Patent No. 5422248  
38 50 16.6 16 1 US-08-366-591-15 Sequence 15, Appl  
39 49.5 16.4 33 1 US-08-237-716-11 Sequence 11, Appl  
40 49.5 16.4 40 1 US-08-099-354-11 Sequence 1, Appl  
41 49.5 16.4 40 2 US-08-288-059-7 Sequence 7, Appl  
42 49.5 16.4 41 1 US-08-361-920-13 Sequence 13, Appl  
43 49.5 16.4 41 1 US-08-479-939-13 Sequence 13, Appl  
44 49.5 16.4 41 1 US-08-483-432-13 Sequence 13, Appl  
45 49 16.2 16 1 US-08-471-033-47 Sequence 47, Appl

## ALIGNMENTS

RESULT 1  
US-08-221-583-56  
; Sequence 56, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 82  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; FILING NUMBER: US/08/221.583  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-221-583-56

Query Match 27.5% Score 83 DB 1 Length 15;  
Best Local Similarity 100.0% Pred. No. 0.0037;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 HLPQPVSTRSQHTQ 29

Db 1 HLPQPVSTRSQHTQ 15

RESULT 2

PCT-US95-04018-56  
; Sequence 56, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-56

Query Match 27.5%; Score 83; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0037;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 HLPQPVSTRSQHTQP 29  
Db 1 HLPQPVSTRSQHTQP 15

RESULT 3  
US-08-221-583-58  
; Sequence 58, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania

; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-58

Query Match 27.2%; Score 82; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SQHTOPTPEPSTAPS 38  
Db 1 SQHTOPTPEPSTAPS 15

RESULT 4  
PCT-US95-04018-58  
; Sequence 58, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04018-58

Query Match 27.2%; Score 82; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SQTQPTPEPSTAPS 38  
|||||  
DB 1 SQTQPTPEPSTAPS 15

RESULT 5  
US-08-221-583-57  
Sequence 57, Application US/08221583  
Patent No. 5486595  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,583  
FILING DATE:  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0185  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-221-583-57

Query Match 26.8%; Score 81; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 STRSQHTQPTPEPST 35  
|||||  
DB 1 STRSQHTQPTPEPST 15

RESULT 6  
US-08-221-583-62  
Sequence 62, Application US/08221583  
Patent No. 5486595  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,583  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0185  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-221-583-62

Query Match 26.8%; Score 81; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 APSTSFLPLPMGSPSP 50  
|||||  
DB 1 APSTSFLPLPMGSPSP 15

RESULT 7  
PCT-US95-04018-57  
Sequence 57, Application PC/TUS9504018  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
APPLICANT: Kruszynski, Marian  
APPLICANT: Mervic, Miljenko  
APPLICANT: Weber, Robert W.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: Norris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-57

Query Match 26.8%; Score 81; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 STRSQHTQTPPEPST 35  
Db 1 STRSQHTQTPPEPST 15

RESULT 8  
PCT-US95-04018-62  
; Sequence 62, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994

; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-62

Query Match 26.8%; Score 81; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 APSTSFLLPMGPSP 50  
Db 1 APSTSFLLPMGPSP 15

RESULT 9  
US-08-221-583-55  
; Sequence 55, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25.mdtctMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-55

Query Match 26.5%; Score 80; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 GAVHLPQPVSTRSQH 26  
Db 1 GAVHLPQPVSTRSQH 15



RESULT 10  
US-08-221-583-59  
; Sequence 59, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcm0d.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-59

Query Match 26.5%: Score 80; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 TQTPPEPSTAPSTSF 41  
Db 1 TQTPPEPSTAPSTSF 15  
RESULT 11  
PCT-US95-04018-55  
; Sequence 55, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-55

Query Match 26.5%: Score 80; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 GAVHLPQPVSTRSQH 26  
Db 1 GAVHLPQPVSTRSQH 15

RESULT 12  
PCT-US95-04018-59  
; Sequence 59, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581

; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US95-04018-59

Query Match 26.5%; Score 80; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TQTPPEPSTAPSTSF 41  
Db 1 TQTPPEPSTAPSTSF 15

RESULT 13  
US-08-221-583-61  
; Sequence 61, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-221-583-61

Query Match 26.2%; Score 79; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0098;  
Matches, 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 PSTAPSTSFLLPMGP 47  
Db 1 PSTAPSTSFLLPMGP 15

RESULT 14  
PCT-US95-04018-61  
; Sequence 61, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US95-04018-61

Query Match 26.2%; Score 79; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0098;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 PSTAPSTSFLLPMGP 47  
Db 1 PSTAPSTSFLLPMGP 15

RESULT 15  
US-08-221-583-60  
; Sequence 60, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris

STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,583  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0185  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-221-583-60

Query Match 25.8%; Score 78; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 30 TPEPSTAPSTSFLLP 44  
|||||  
Db 1 TPEPSTAPSTSFLLP 15

Search completed: February 12, 2002, 13:00:17  
Job time: 112 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:45:58 ; Search time 19.28 Seconds  
(without alignments)  
1821.394 Million cell updates/sec

Title: US-09-800-909-2  
Perfect score: 2468  
Sequence: 1 MAPVAWAAALAVGLELWAAA.....GSTEERPLPLGVDPDAGMKPS 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2468	100.0	461	1 A35356	tumor necrosis fac
2	1512	61.3	474	2 B38634	tumor necrosis fac
3	1477	59.8	459	2 I48854	gene murine tumour
4	381.5	15.5	349	2 D72175	G2R protein - vari
5	379.5	15.4	348	2 T28623	hypothetical prote
6	379.5	15.4	349	2 D36858	gene G4R protein -
7	360.5	14.6	326	1 G0V2ML	T2 protein - myxom
8	341.5	13.8	435	2 I54182	tumor necrosis fac
9	328	13.3	325	2 B43692	T2 protein - rabbi
10	317.5	12.9	277	2 A60771	B-cell activation
11	238.5	10.5	305	2 A46476	B cell-associated
12	251.5	10.2	416	1 JN0006	nerve growth facto
13	241	9.8	595	2 A42086	CD30 antigen precu
14	229.5	9.3	271	2 S12783	OX40 antigen precu
15	223.5	9.1	272	2 I48700	gene ox40 protein
16	219.5	8.9	493	2 J55486	membrane glycoprot
17	217	8.8	277	2 I37552	OX40 homolog - hum
18	210	8.5	427	1 GQHUN	nerve growth facto
19	198	8.0	461	2 J43302	tumor necrosis fac
20	197	8.0	256	2 B32393	T-cell antigen 4-1
21	197	8.0	425	1 A36431	nerve growth facto
22	191.5	7.8	454	1 G0MST1	tumor necrosis fac
23	185.5	7.5	461	1 G0RTT1	tumor necrosis fac
24	185	7.5	255	2 T38426	lymphocyte activat
25	183.5	7.4	455	1 G0HUT1	tumor necrosis fac
26	166.5	6.7	1367	1 S48478	glucan 1,4-alpha-g
27	165	6.7	770	2 T51024	related to C2H2 zi
28	161.5	6.5	1203	2 T17415	mycellial surface a
29	161	6.5	1372	2 T25933	hypothetical prote

30 159.5 6.5 2232 2 T34434  
31 158 6.4 1274 2 T42017  
32 156.5 6.3 383 2 T46707  
33 154.5 6.3 1428 2 T08852  
34 153.5 6.2 3507 2 T34513  
35 150 6.1 801 2 T29018  
36 148.5 6.0 1032 2 T34433  
37 148 6.0 327 2 A46484  
38 147 6.0 438 2 T31889  
39 147 6.0 600 2 S07638  
40 145 5.9 742 2 I37225  
41 144 5.8 314 2 I37383  
42 143.5 5.8 324 2 JC2395  
43 143 5.8 534 2 T39903  
44 143 5.8 3942 2 T42730  
45 142 5.8 1151 2 T18535

## ALIGNMENTS

RESULT 1

A35356  
tumor necrosis factor receptor 2 precursor [validated] - human  
N:Alternate names: 75K tumor necrosis factor receptor; TNF receptor type 2  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 08-Dec-2000  
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094  
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, Science 248, 1019-1023, 1990  
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a  
A:Reference number: A35356; MUID:90260639  
A:Accession: A35356  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-461 <SNL>  
A:Cross-references: GB:M32315; NID:gl189185; PIDN:AAAS9929.1; PID:gl189186  
R:Kohno, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990  
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc  
A:Reference number: A36475; MUID:91045991  
A:Accession: A36475  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-195, 'R', 197-461 <KOH>  
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758  
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, Cytokine 2, 231-237, 1990  
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular  
A:Reference number: A48416; MUID:91370690  
A:Accession: A48416  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 23-461 <DEM>  
A:Cross-references: GB:S63368; NID:g235648; PIDN:AAAB19824.1; PID:g235649  
A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBI:63371)  
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990  
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons  
A:Reference number: A36007; MUID:90349572  
A:Accession: A36007  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>  
A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752  
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, J. Biol. Chem. 265, 20131-20138, 1990  
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor  
A:Reference number: A23666; MUID:91056048  
A:Accession: A23666  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 23-40; 65-69; 136-141; 300-306 <LOE>

hypothetical prote  
cysteine rich prot  
proteophosphoglyca  
lustrin A - Califo  
hypothetical prote  
hypothetical prote  
hypothetical prote  
apoptosis-mediati  
hypothetical prote  
spore coat protein  
leucocyte antigen  
FAS soluble protei  
FAS antigen precu  
serine-rich protei  
Bassoon protein -  
high molecular mas

R:Engelmann, H.; Novick, D.; Wallach, D.  
 J. Biol. Chem. 265, 1531-1536, 1990  
 A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evidence  
 A:Reference number: A35010; MUID:90110215  
 A:Accession: B35010  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 27-31 <ENG>  
 R:Kuhnert, P.; Kemper, O.; Wallach, D.  
 Gene 150, 381-386, 1994  
 A:Title: Cloning, sequencing and partial functional characterization of the 5' region of  
 A:Reference number: I38094; MUID:95121934  
 A:Accession: I38094  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-37 <RES>  
 A:Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701  
 C:Genetics:  
 A:Gene: GDB:TNFR2  
 A:Cross-references: GDB:I25914; OMIM:191191  
 A:Map position: lp36.2-lp36.2  
 A:Introns: 26/3  
 A>Note: the list of introns is incomplete  
 C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-416/Product: tumor necrosis factor receptor 2 #status experimental <MAT>  
 F:40-76/Domain: NGF receptor repeat homology <NG1>  
 F:78-119/Domain: NGF receptor repeat homology <NG2>  
 F:120-162/Domain: NGF receptor repeat homology <NG3>  
 F:164-201/Domain: NGF receptor repeat homology <NG4>  
 F:262-279/Domain: transmembrane #status predicted <TMN>  
 F:280-461/Domain: intracellular #status predicted <INT>  
 F:171.193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 2468; DB 1; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-134;  
 Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLEHYDQTAQMCCSKSPG 60  
 |||||  
 DB 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLEHYDQTAQMCCSKSPG 60  
 QY 61 QHAKVFCTKTSDFVCDSCEDSTYTLQNNWVPECLSCGSRSSDQVETQACTREONRICTC 120  
 |||||  
 DB 61 QHAKVFCTKTSDFVCDSCEDSTYTLQNNWVPECLSCGSRSSDQVETQACTREONRICTC 120  
 QY 121 RPYWCALSKQECRLCAPLRCRPGFGVARPGETSDVYCKPCAPGTFNTTSSSTDICR 180  
 |||||  
 DB 121 RPYWCALSKQECRLCAPLRCRPGFGVARPGETSDVYCKPCAPGTFNTTSSSTDICR 180  
 QY 181 PHQICNVAIPGNASMDVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
 |||||  
 DB 181 PHQICNVAIPGNASMDVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
 QY 241 FLPLPMGSPSPAEGSTGDFALPVLGVGTALGLLITGVNVCMTQVKKPLCLQREAKV 300  
 |||||  
 DB 241 FLPLPMGSPSPAEGSTGDFALPVLGVGTALGLLITGVNVCMTQVKKPLCLQREAKV 300  
 QY 301 PHLPADKARGTQPEQOHLITAPSSSSSSLESSASALDRAPTRNQOPAGVEASGAGE 360  
 |||||  
 DB 301 PHLPADKARGTQPEQOHLITAPSSSSSSLESSASALDRAPTRNQOPAGVEASGAGE 360  
 QY 361 ARASTGSSDSSPGGHGTQVNVTCIVNVCSSDSSHSSQCSSOASSTMGDTDSSSPKDEQ 420  
 |||||  
 DB 361 ARASTGSSDSSPGGHGTQVNVTCIVNVCSSDSSHSSQCSSOASSTMGDTDSSSPKDEQ 420  
 QY 421 VPESKEECAPFRSQLETPTETLLGSTEKPLPLGVDPAGMKPS 461  
 |||||  
 DB 421 VPESKEECAPFRSQLETPTETLLGSTEKPLPLGVDPAGMKPS 461

## RESULT 2

B38634  
 tumor necrosis factor receptor type 2 precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
 C:Accession: B38634; A40254; S54816  
 R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chan,  
 Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991  
 A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis facto  
 A:Reference number: A38634; MUID:91187885  
 A:Accession: B38634  
 A:Molecule type: mRNA  
 A:Residues: 1-474 <LEM>  
 R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J  
 Mol. Cell. Biol. 11, 3020-3026, 1991  
 A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f  
 A:Reference number: A40254; MUID:91246168  
 A:Molecule type: mRNA  
 A:Residues: 1-474 <GOO>  
 A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828  
 R:Kisssonerhis, M.; Fellowes, R.; Feldmann, M.; Chernajovsky, Y.  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Characterization of the promoter region of the murine p75-TNF receptor  
 A:Reference number: S54816  
 A:Accession: S54816  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-22 <KIS>  
 C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
 C:Keywords: cytokine receptor; transmembrane protein  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>  
 F:40-77/Domain: NGF receptor repeat homology <NG1>  
 F:79-120/Domain: NGF receptor repeat homology <NG2>  
 F:166-203/Domain: NGF receptor repeat homology <NG4>  
 Query Match 61.3%; Score 1512; DB 2; Length 474;  
 Best Local Similarity 63.0%; Pred. No. 1.7e-79;  
 Matches 293; Conservative 50; Mismatches 114; Indels 8; Gaps 6;  
 QY 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRL-REYDQTAQMCCSKSP 59  
 |||||  
 DB 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRL-REYDQTAQMCCSKSP 60  
 QY 60 QHAKVFCTKTSDFVCDSCEDSTYTLQNNWVPECLSCGSRSSDQVETQACTREONRICT 119  
 |||||  
 DB 61 QHAKVFCTKTSDFVCDSCEDSTYTLQNNWVPECLSCGSRSSDQVETQACTREONRICT 120  
 QY 120 RPYWCALSKQECRLCAPLRCRPGFGVARPGETSDVYCKPCAPGTFNTTSSSTDICR 178  
 |||||  
 DB 121 RPYWCALSKQECRLCAPLRCRPGFGVARPGETSDVYCKPCAPGTFNTTSSSTDICR 180  
 QY 179 PHQICNVAIPGNASMDVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 238  
 |||||  
 DB 181 PHQICNVAIPGNASMDVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 239  
 QY 239 TSFLPMGSPSPAEGST-GDFALPVLGVGTALGLLITGVNVCMTQVKKPLCLQRE 297  
 |||||  
 DB 240 TSFLPMGSPSPAEGST-GDFALPVLGVGTALGLLITGVNVCMTQVKKPLCLQRE 298  
 QY 298 AKVPHLPADKARGTQPEQOHLITAPSSSSSSLESSASALDRAPTRNQOPAGV-EAS 356  
 |||||  
 DB 299 AKVPHLPADKARGTQPEQOHLITAPSSSSSSLESSASALDRAPTRNQOPAGV-EAS 358  
 QY 357 GAGARASTGSSDSSPGGHGTQVNVTCIVNVCSSDSSHSSQCSSOASSTMGDTDSSSPESP 416  
 |||||  
 DB 359 GAGARASTGSSDSSPGGHGTQVNVTCIVNVCSSDSSHSSQCSSOASSTMGDTDSSSPESP 418  
 QY 417 KQEQVPSKEECAPFRSQLETPTETLLGSTEKPLPLGVDPAGMKPS 461

Db 148854 KDEQVPSFQEECPSPQCTETTEL--QSHEKPLPLGVDPDMGPKS 461  
: |||||: ||| : || ||| : ||||| ||||| |||||  
RESULT 3  
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I48854  
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.  
Mamm. Genome 5, 726-727, 1994  
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.  
A:Reference number: I48854; MUID:95178848  
A:Accession: I48854  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-459 <RES>  
A:Cross-references: EMBL:X76401; NID:g433830; PIDN:CAA53981.1; PID:g433831  
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
F:151-188/Domain: NGF receptor repeat homology <NGF>  
Query Match 59.8%; Score 1477; DB 2; Length 459;  
Best Local Similarity 63.8%; Pred. No. 1.6e-77;  
Matches 287; Conservative 47; Mismatches 108; Indels 8; Gaps 6;  
Qy 16 LWAAHAHALPAQVAETPYAPEPGSTCL-REYYDQTAQMCCSCSPGQAHKVFCTKTSDTV 74  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 LWATGHTVPAQVLLTPYKPEPGYEQIQEYVDRAKMCACKPPGQYVYKFCNKTSDIV 60  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 75 CDSCESTYTLWNVWVPECLSGSRCSDDQVETQACTRQNRICRCPGWYCALSKQEG- 133  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 61 CADCEASMYTQWQFRCLSCSSCSDDQVETRACTKQNRVCAACEAGRYCALKTHSGS 120  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 134 CRLCAPLKKCRPGFGVARGTETSDVCKPCAPGTFSTNTSTDICRPHQICNVVAIPGN 193  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 121 CQCRLSKCGPGFGVASSRPNAGNVLCACAPGTFSDTSTSDVCRPHRISILAIIPGN 180  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 194 ASMDAVCFSTPTSRMAGFALPQVPVSTRSQHTQPTPEPSTAPSTSFLLPMGPPSPAE 253  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 181 ASTDAVCAPEFTLSAIPRTLYVSPETRSQPLQEPGSPQTP--SILTSIGSTPIEQ 238  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 254 ST-GDFALPLVGLVGTALGLLIIGVNCVIMTVQKKPLCLQREAKVPHLPADKARGTQ 312  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 239 STKGGISLPIGLVGTSLGLLMLGLVNCFILVQRKKPKSCLQORDAKVPHVPDEKSQDAV 298  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 313 GPEQOHLITAPSSSSSSLESSASALDRAPTRNQPAQGV-EASGAGEARASTGSSDSS 371  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 299 GLEQOHLITAPSSSSSSLESSASAGDRAPPGGHPQARVMAEQSQEARASSRISDSS 358  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 372 PGGHGTQVNVTCIVNVCSSDHSSCSQASSTMGDTDSSPSKDEQVPFSKKECAFR 431  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 359 HSGHGTQVNVTCIVNVCSSDHSSCSQASATVGDPAKPSAKDEQVPFSQEECPSQ 418  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 432 SQLETPETLLGSTEKPLPLGVDPDMGPKS 461  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 419 SPYETTEL--QSHEKPLPLGVDPDMGPKS 446  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
RESULT 4  
G2R protein - variola minor virus (strain Garcia-1966)  
C:Species: variola minor virus  
C:Date: 24-Nov-1999 #sequence\_revision 24-Nov-1999 #text\_change 20-Jun-2000  
C:Accession: D72175  
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopad  
submitted to GenBank, March 1998  
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor  
A:Reference number: A72150  
A:Accession: D72175  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-349 <SHC>  
A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54798.1; PID:g5830759  
A:Experimental source: strain Garcia-1966  
C:Genetics:  
A:Gene: G2R  
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology  
Query Match 15.5%; Score 381.5; DB 2; Length 349;  
Best Local Similarity 38.1%; Pred. No. 4.1e-15;  
Matches 80; Conservative 23; Mismatches 86; Indels 21; Gaps 7;  
Qy 31 PYAPEPGSTCLREYYDQTAQMCCSKSPGQAHKVFCTKTSDDTCDSCEDSTYTLWNVW 90  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 24 PYTP-PNGKCKDTEY--KRNLCLCLSCPPGTYASRLCDSKNTQCTPCGSGTFTSRNNHL 80  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 91 PECLSGSRCSDDQVETQACTRQNRICRCPGWYCALSKQEGRLCAPLKKCRPGFGVA 150  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 81 PACLSGCRGNSQVETRSCNTTHNRICECSPGYICLLKSGSGCKACVQSQTKCGIGYGV 140  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 151 RPTETSDVVKPCAPGTFSTNTSTDICRPHQICNVVAIPGNA--SMDAVCT-----ST 203  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 141 -GHTSVGDVICSPCGFGTYSVTSSDKCEP-----VPNNTFNYIDVEITLVPVNDT 191  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 204 SPTRSMAPGAVH--LPQPVSTRSQHTQTP 231  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 192 SCTRTTTTGLSEILTSELITMTNMHTDNP 221  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
RESULT 5  
T28623  
hypothetical protein G2R - variola major virus  
C:Species: variola major virus  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T28623  
R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Au  
Nature 366, 748-751, 1993  
A:Title: Potential virulence determinants in terminal regions of variola smallpox vir  
A:Reference number: Z20488; MUID:94088747  
A:Accession: T28623  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-348 <MAS>  
A:Cross-references: EMBL:L22579; NID:g623595; PIDN:AAA60933.1; PID:g439102  
A:Experimental source: strain Bangladesh 1975  
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology  
Query Match 15.4%; Score 379.5; DB 2; Length 348;  
Best Local Similarity 37.6%; Pred. No. 5.3e-15;  
Matches 79; Conservative 24; Mismatches 86; Indels 21; Gaps 7;  
Qy 31 PYAPEPGSTCLREYYDQTAQMCCSKSPGQAHKVFCTKTSDDTCDSCEDSTYTLWNVW 90  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 23 PYTP-PNGKCKDTEY--KRNLCLCLSCPPGTYASRLCDSKNTQCTPCGSGTFTSRNNHL 79  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 91 PECLSGSRCSDDQVETQACTRQNRICRCPGWYCALSKQEGRLCAPLKKCRPGFGVA 150  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 80 PACLSGCRGNSQVETRSCNTTHNRICECSPGYICLLKSGSGCKACVQSQTKCGIGYGV 139  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 151 RPTETSDVVKPCAPGTFSTNTSTDICRPHQICNVVAIPGNA--SMDAVCT-----ST 203  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 140 -GHTSVGDVICSPCGFGTYSHTVSSADKCEP-----VPNNTFNYIDVEITLVPVNDT 190  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 204 SPTRSMAPGAVH--LPQPVSTRSQHTQTP 231  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 191 SCTRTTTTGLSEILTSELITMTNMHTDNP 220  
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
RESULT 6  
D36858  
gene G4R protein - variola virus  
N:Alternate names: B28R protein (COP)

```

QY      215 HLPQP-----VST-----RSQHTQTPPEPSTAPSTSFLL---PMGFSP-PAEGS----- 254
          ||      ::||      ||      ||      ||      ||      ||      ||

```

Db 231 LLPLAFFLLLATVFCIMKRS-----HPSLCKLGLSLLKRRPQGGPNPVAGSWPEPPKA 283  
QY 255 -----TGDFALPVLGVITAGLLIIGVNVCMVMTQVKKKPLCLQREAKV 300  
Db 284 HPYFDPVLQPLPISGDVS-PVS--TGLPAAPVLEAGVPQ-----QQSPDLDTRE--- 330  
QY 301 PHL-PADNARGTQGPQEQHLLITAPSSSSSSLESSASALDRAPTRNQPOAPG 352  
Db 331 POLEPGEQSOVAHGTNGIHV-----TGGSMITGNIYIYNGPVLLGGPPGPG 376

RESULT 9  
B43692  
T2 protein - rabbit fibroma virus  
C:Species: rabbit fibroma virus, Shope fibroma virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: B43692  
R:Upton, C.; DeLange, A.M.; McFadden, G.  
Virolgy 160, 20-30, 1987  
A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric  
A:Reference number: A43692; MUID:87321103  
A:Accession: B43692  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-325 <UPT>  
A:Cross-references: GB:M17433  
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homolog  
F:64-105/Domain: NGF receptor repeat homolog <NG2>  
F:106-147/Domain: NGF receptor repeat homolog <NG3>

Query Match 13.3%; Score 328; DB 2; Length 325;  
Best Local Similarity 36.9%; Pred. No. 4.3e-12;  
Matches 66; Conservative 28; Mismatches 71; Indels 14; Gaps 6;  
QY 31 PYAPEPGSTCLREYDYDPAQMCCKSCPGQAKVFCVTKTSDTVCDSCEDSTVQLNWW 90  
Db 20 PYSSNQK-CGGHDY--EKDGLCCASCHPEFYASRLCGGSGNFVCSPECDGTFTASTNHA 76  
QY 91 PECLSCGSSCDSDQVETQACTREQNRICPCRPWYCALSKQEGCRCLAPLRCRPGFGVA 150  
Db 77 PACVSCRGPTGHLSESPQCDRTHRVNCSTGNYCLLKGQNGCRICAPQTKCPAGYGV 136  
QY 151 RPTETSDVWCKPAPGTFSTNTSDTICRPHQICNVAI-----PGNASMDAVCTSTS 204  
Db 137 -GHTRAGDTLCEKPPHTYSDLSPTERCSTGTS--FNYISVGFNLYPVN---ETSCITTA 189

RESULT 10  
A60771  
B-cell activation protein CD40 precursor - human  
N:Alternate names: B-cell surface antigen Bp50  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Feb-1994 #text\_change 21-Jul-2000  
C:Accession: S04460; A60771  
R:Stamenkovic, I.; Clark, E.A.; Seed, B.  
EMBO J. 8, 1403-1410, 1989  
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor receptor  
A:Reference number: S04460; MUID:89356608  
A:Accession: S04460  
A:Molecule type: mRNA  
A:Residues: 1-277 <STA>  
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CRAA43045.1; PID:g29851  
R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.  
J. Immunol. 142, 562-567, 1989  
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-like  
A:Reference number: A60771; MUID:89093941  
A:Accession: A60771  
A:Molecule type: protein  
A:Residues: 21-50 <BRA>  
A:Experimental source: Burkitt lymphoma cell line Raji  
C:Genetics:  
A:Gene: GDB:CD40

A:Cross-references: GDB:215268; OMIM:109535  
A:Map position: 20q12-20q13.2  
C:Superfamily: CD27 antigen; NGF receptor repeat homolog  
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>  
F:21-193/Domain: extracellular #status predicted <EXT>  
F:194-215/Domain: transmembrane #status predicted <TM>  
F:216-277/Domain: intracellular #status predicted <CYT>  
F:153,180/Binding site: carbohydrate (Asn) #status predicted

Query Match 12.9%; Score 317.5; DB 2; Length 277;  
Best Local Similarity 26.7%; Pred. No. 1.5e-11;  
Matches 92; Conservative 32; Mismatches 124; Indels 97; Gaps 11;  
QY 23 LPAQVA-----FTPYAPEPGSTCLREYDYDPAQMCCKSCPGQAKVFCVTKTSDTVCD 77  
Db 4 LPLQCVLWGCLLTAVHPEPTACREKQVLYNS--QCCSLCQPGOKLVSDCTETETECLP 61  
QY 78 CEDSTYQLNWNVPEC-----LSCGSRCSDDQVETQACTREQNRICTCRPGWYCALSK 130  
Db 62 CGESEFLDTWNRETHCHQHKYCDPNLGRVQ-----OKGTSETDTICTCEEGRHCT--- 112  
QY 131 QEGCRLCAPLRCRPGFGVAPGTETSDVCKPCAPGTFSNTTSDTICRPHQICN---- 186  
Db 113 SEACESCVLHRSKSPGFGVKQIATGVSDTICEPCPVGFFSNVSSAFKCHPMTSCETKDL 172  
QY 187 VVAIPGNASMDAVCTSTSPTRSMAPGAVHLPPQVSTRSQHTQPTPEPSTAPSTSELLPMG 246  
Db 173 VVQAGTNTKTDVCGPQDRLEAL----- 195  
QY 247 PSPPAEGSTGDFALPVLGVITAGLLIIGVNVCMVMTQVKKKPLCLQREAKVPH---- 302  
Db 196 -----VVPIIFGILEFALLVL-----VFIKKVAKKP-----TNKAPHPKE 232  
QY 303 -----LPAD-KARTQGPQEQHLLITAPSSSSSSLESSASALDRR 341  
Db 233 PQEINFDDLPGSNTAAPVQETHGCPVQEDGKESRISVQERQ 277

RESULT 11  
A6476  
B cell-associated surface molecule CD40, long splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Nov-2000  
C:Accession: A46476; A46515  
R:Torres, R.M.; Clark, E.A.  
J. Immunol. 148, 620-626, 1992  
A:Title: Differential increase of an alternatively polyadenylated mRNA species of mur  
A:Reference number: A46476; MUID:92105763  
A:Accession: A46476  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-305 <TOR>  
A:Cross-references: GB:M83312; NID:g1553058  
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)  
A:Note: this translation is not annotated in GenBank entry MUSCD40A, release 113.0  
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockay  
J. Immunol. 149, 3921-3926, 1992  
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.  
A:Reference number: A46515; MUID:93094586  
A:Accession: A46515  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-287, 'LV' <GR>  
A:Cross-references: GB:M83312; NID:g1553058; PIDN:AAB08705.1; PID:g1553059; GB:M94126  
A:Experimental source: BALB/c, liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)  
A:Note: For an alternative splice form, see PIR:A46515.  
C:Comment: For an alternative splice form, see PIR:A46476.  
C:Superfamily: CD27 antigen; NGF receptor repeat homolog  
C:Keywords: alternative splicing; transmembrane protein



F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 10.5%; Score 258.5; DB 2; Length 305;  
Best Local Similarity 22.8%; Pred. No. 3.8e-08;  
Matches 87; Conservative 37; Mismatches 142; Indels 115; Gaps 14;  
QY 5 AVNALAVGLLEWAAHALPAQVAFTPYAPEPGSTCLREYYDQTAQMCCKSPGQHAQ 64  
DB 9 ALMGCLLTAVHLGCV-----TCSDKQYLHD--GQCCDLCPGSRLL 48  
QY 65 VFTKTSDDTCDSCEDSTYTQLNWVPECLSCGSRCSDDQ---VETQACTREQNRICR 121  
DB 49 SHCTALEKTOCHPCDSEGEFAQWNRIRCHO--HRHCEPNOGLRVKKEG--TAESDTVCCK 106  
QY 122 PGWYCALSKGEGRLCAPLKRCRFGVARGTETSDVCKPCAPGTFNNTSTSDICRP 181  
DB 107 EGQCHT---SKDCACAAHQHPCIPGFGVMEMATETTDVCHPCPGVFGFSNOSSLFKCY 163  
QY 182 HQICN-----VVAIPGNASMDAVCTSTPTSRMAPGAVHLPOPVSTRSQHTOPTPEPSTAP 237  
DB 164 WTSCEDKNLEVLQKGTQTNVICLKSRRAL----- 195  
QY 238 STSFLLPMGSPPAEGTGFDPALVGLVGTALGLLIIGVWNCVIMTQVKKPLCLQRE 297  
DB 196 -----LVIPVMGILITIFGVFL-----YIKKVKKP---KDN 225  
QY 298 AKVPHLPADK-----ARTQGPQEOHLLITAPSSSSSSLESSASALDR----- 340  
DB 226 EMPL--PAARRQDQEMEDYVGHNTAAPVQETHGCPVQEDGKESRISVOERQVDSI 283  
QY 341 --RAPTRNPQAPGVASGAG 359  
DB 284 ALRPPGLN----PCTATGGDG 300

## RESULT 12

JN0006  
N:Alternate names: NGF receptor  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
R:Accession: JN0006; A60504  
R:Large, T.H.; Weskamp, G.; Helder, J.C.; Radeke, M.J.; Misko, T.P.; Shooter, E.M.; Reid  
Neuron 2, 1123-1134, 1989  
A:Title: Structure and developmental expression of the nerve growth factor receptor in t  
A:Reference number: JN0006; MUID:90186579  
A:Accession: JN0006  
A:Molecule type: mRNA  
A:Residues: 1-416 <LAR>  
A:Experimental source: embryonic chick brain  
R:Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.  
Dev. Biol. 137, 287-304, 1990  
A:Title: Structure and developmental expression of the chicken NGF receptor.  
A:Reference number: A60504; MUID:90152140  
A:Accession: A60504  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 21-35, 'Y', 37-172, 'K', 174-275, 'S', 277-395, 'R', 397-416 <HEU>  
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma  
C:Comment: The cysteine-rich region of the extracellular domain may form part or all of  
C:Comment: This protein is thought to form a high-affinity receptor when it associates  
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology  
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; t  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-416/Product: nerve growth factor receptor #status predicted <MAT>  
F:21-239/Domain: extracellular #status predicted <EXT>  
F:24-57/Domain: NGF receptor repeat homology <NG1>  
F:59-100/Domain: NGF receptor repeat homology <NG2>  
F:101-139/Domain: NGF receptor repeat homology <NG3>  
F:141-181/Domain: NGF receptor repeat homology <NG4>  
F:189-237/Region: serine/threonine-rich  
F:240-261/Domain: transmembrane #status predicted <MEM>

F:262-416/Domain: intracellular #status predicted <INT>  
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.2%; Score 251.5; DB 1; Length 416;  
Best Local Similarity 25.1%; Pred. No. 1.3e-07;  
Matches 89; Conservative 48; Mismatches 164; Indels 53; Gaps 16;  
QY 23 LPAQVAFTPYAPEPGS--TCRLREYYDQTAQMCCKSPGQHAQVFTKTSDDTCDSCED 80  
DB 5 VFLLLLLLPAQVAGTWSGKELTKMY--TTSGECKACNLGEGVVQPC-GVNTQVCEPCLD 61  
QY 81 S-TYQLNWNWPECLSCGSRCSDDQVETQACTREQNRICRTPGWCYCALSKQEGRLCAP 139  
DB 62 SVTYSDTVSATPEKPC-TQCGLHSMSPACVESDADVCAVGYF---QDELSGSCKE 116  
QY 140 LRKCRPGCVARPGTETSDVCKPCAPGTFNNTSTSDICRPHQICNVVAIPGNASMDAV 199  
DB 117 CSICEVGFGLMFPKCRDSQPTVCECEGTFSEANFVDPCLPCTICE-----ENEMVKE 171  
QY 200 CTSTSTPT--RSMAP--GAVHLPOPVSTRSQHTOPTPEP-----STAPSTSFLLPM 245  
DB 172 CTATSDAECRLDHPWTHTPSLAGSDS-----PEPITRDPENTEGMATTLADIYTVM 225  
QY 246 GPSPP-AGESTGDFALPVGLIIVGTALGLLIIGVWNCVIMTQVKKPLCLQREAKVPHLP 304  
DB 236 GSSQPVWSRGTDNLIPVCSI---LAAVVGLVAYIAF---KRWNSCKQKQGANRP 278  
QY 305 ADKARGTGPQEOHLLITAPSSSSSSLESSASALDRRAPTRNPQAPGVASGA 358  
DB 279 VNQ---TPSPEKEL-----HSDSGISVDSQSLHDQPPNQSTQGPAPKGDGS 323

## RESULT 13

A42086  
CD30 antigen precursor - human  
N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jul-2000  
C:Accession: A42086  
R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.  
Cell 68, 421-427, 1992  
A:Title: Molecular cloning and expression of a new member of the nerve growth factor  
A:Reference number: A42086; MUID:92154659  
A:Accession: A42086  
A:Molecule type: mRNA  
A:Residues: 1-595 <DUR>  
A:Cross-references: GB:M83554; NID:g180095; PIDN:AAA51947.1; PID:g180096  
A:Experimental source: HUT-102 cell line  
A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:82090)  
C:Genetics:  
A:Gene: GDB:CD30; D1S166E  
A:Cross-references: GDB:I31547; OMIM:153243  
A:Map position: lp36-lp36  
C:Superfamily: NGF receptor repeat homology  
C:Keywords: glycoprotein; growth factor receptor; transmembrane protein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-38/Domain: extracellular #status predicted <EXT>  
F:384-407/Domain: transmembrane #status predicted <TMM>  
F:408-595/Domain: intracellular #status predicted <CYT>  
F:101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 241; DB 2; Length 595;  
Best Local Similarity 22.0%; Pred. No. 7.1e-07;  
Matches 113; Conservative 41; Mismatches 183; Indels 176; Gaps 17;  
QY 11 AVGLELWAAHALPAQVAFTPYAPEPGSFC--RLREYYDQTAQMCCKSPGQHAQVFT 68  
DB 7 ALGLLFLGALRAFPQDRPFE-----DTCHGNPNSHYDKAVRRCRCYPMGLFPPTQQCP 59  
QY 69 KTSDDTVCDSCEDSTYTQLNWVPECLSCGSRCSDDQVETQACTREQNRICRTPGWCYCAL 128

```
Db 60 QRPDCRKOCEPDYLL---DEADRCACVTCGRDDLVEKTPCAWNSSRVCECRPGMFCST 116
Qy 129 SKQECRLCAPLRKCRPGFVGARPGTETSDVCKPCAPG-----167
Db 117 SAVNSCARCFFHVSVPAGMIVKPEPGTAQKNTVCEPASPGVPACASPENCKEPPSSGTIPQ 176
Qy 168 -----TFSNTT-----173
Db 177 AKPTVPSTATSSASTWVPVRGGTRLAQEAASKLTRAPDSPSSVGRPSDDGLSPTQPCPEG 236
Qy 174 -----SSTDCRPHQICNVVAIPG 192
Db 237 SGDCRKOCEPDYLLDEAGRCTACVSCSRDDLVEKTPCAWNSSRTCECRGMICATSATNS 296
Qy 193 NASMDAVCTSTSPTRSMAPGAVHLFQ-----PVSTRSQHTQTPPE-----PST 235
Db 297 -----CARCVYPICAAETVTKPDMAEKDTTFEAPPLGTQPD-CNPTPENGEPAST 348
Qy 236 APSTSFLL-----LPMGPSP-PEGSGDFALPVLGVGTALGLLIIGVVNCVIMTQ 286
Db 349 SPTQSLLDVDSQASKTLPITPSAPVALSSSTGKPVLDAGPVLFWILVLVVVVGSSAFLLCH 408
Qy 287 VKKKPLCLQREAKVPHL---PADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPT 344
Db 409 ---RRACRKRIKQLHLCPVQ---TSQPKLE-LVDSRPRSSTQLRSGASVTEPVAEE 460
Qy 345 R---NOPQAPGVEASGAGARASTGSSDSPGG 374
Db 461 RGLMSOPLMETCHSVGAAYLE-SLPLODASPAG 492
RESULT 14
SI2783
N:Altenate names: rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: SI2783; S08036
R:Mallett, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocyte
A:Reference number: SI2783; MUID:90214614
A:Accession: SI2783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:957830; PIDN:CAA34897.1; PID:957831
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-271/Product: OX40 antigen #status predicted <MAT>
F:211-235/Domain: transmembrane #status predicted <TM>
Query Match 9.3%; Score 229.5; DB 2; Length 271;
Best Local Similarity 27.1%; Pred. No. 1.5e-06;
Matches 79; Conservative 28; Mismatches 88; Indels 97; Gaps 14;
Qy 6 VW-----AALAVGLELWAAHALPAQVATPYAPEGSGTCRLREYD--QTAQWCCSKCS 58
Db 3 VWQOQPTALLGLLSL-----GVTVLCNKVDTPSGHKCCRECC 42
Qy 59 PGQAKVFCCTKSDTVCDSCEDSTYTLQMNW--VPECLSGSGSCSDQVETQACTREONR 116
Db 43 PGHGVSRCDHTRDTVCHPCETGTFYNEAVNYDTCKOCTQCNHRSGSEL--KQNCPTPTEDT 100
Qy 117 ICTCRPGWYCALSKQEGCRCLAPLRKCRPGFVGARPGTETS---DVVCKPCAPGTFESNTT 173
Db 101 VCOCREG-----TQPRQDSSHLKGVDCVPCPGHFS--P 132
Qy 174 SSTDCRPHQICNV-----VAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT-- 227
Db 133 GSNQACKPMTNCTLSGKQIRHPASNSLDTVCEDRS-----LLATLLWETQRTTF 181
```

```
Qy 228 OPTPEPSTA-----PSTSFLL-PMGSPSPAEGSTGDFALPVLGLIVGVT 270
Db 182 RPTTVPSTTVWPRISQPLSPPTLVAPEGPA-----FAVILGLGLGLLA 224
RESULT 15
I48700
gene ox40 protein - mouse
N:Alternate names: OX40 antigen
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 11-Jan-2000
C:Accession: I48700; I48334; S34377
R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J
J. Immunol. 151, 5261-5271, 1993
A:Title: Cloning of mouse OX40: a T cell activation marker that may mediate T-B cell
A:Reference number: I48700; MUID:94044750
A:Accession: I48700
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-272 <RES>
A:Cross-references: EMBL:Z21674; NID:9312827; PIDN:CAA79772.1; PID:g312828
R:Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
Eur. J. Immunol. 25, 926-930, 1995
A:Title: Gene structure and chromosomal localization of the mouse homologue of rat OX
A:Reference number: I48334; MUID:95255413
A:Accession: I48334
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-14 'G', 16-272 <RE2>
A:Cross-references: EMBL:X85214; NID:g732818; PIDN:CAA59476.1; PID:g732819
C:Genetics:
A:Gene: OX40
A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1
C:Superfamily: CD27 antigen; NGF receptor repeat homology
Query Match 9.1%; Score 223.5; DB 2; Length 272;
Best Local Similarity 27.0%; Pred. No. 3.3e-06;
Matches 88; Conservative 32; Mismatches 97; Indels 109; Gaps 19;
Qy 6 VW-----AALAVGLELWAAHALPAQVATPYAPEGSGTCRLREYDQTAQWCCSKCSPG 60
Db 3 VWQOQPTALLGLLSLTLGVTARRL-----NCVKHTY--PSGHKCCRECCQ 45
Qy 61 QHAKVFCCTKSDTVCDSCEDSTYTLQMNW--VPECLSGSGSCSDQVETQACTREONR 118
Db 46 HGMVSRCDHTRDTLCHPCETGTFYNEAVNYDTCKOCTQCNHRSGSEL--KQNCPTPTQDTVC 103
Qy 119 TCRPGWYCALSKQEGCRCLAPLRKCRPGFVGARPGTETS---DVVCKPCAPGTFESNTSS 175
Db 104 RCRPG-----TQPRQDSSGYKLGVDVCPGPHFS--PGN 135
Qy 176 TDCRPHQICNV-----VAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQ-PT 230
Db 136 NQACKPMTNCTLSGKQIRHPASNSLDAVCE-----RSL-----LATLLWETQRTPT 181
Qy 231 PEPSTA-----PSTSFLLPMGPSPPA-----EGSGDFALPVLGLIVGVTALGLLIIGVV 279
Db 182 FRPTVSTTVWPRISQPLSPPTLVAPEGPA-----FAVILGLGLGLLAPLTVLL---- 231
Qy 280 NCVIMTQVKKKPLCLQREA-KVPHLP 304
Db 232 -----ALYLLRKAWRLPNTP 246
Search completed: February 12, 2002, 12:48:26
Job time: 148 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:48:04 ; Search time 13.25 Seconds  
(without alignments)  
1275.659 Million cell updates/sec

Title: US-09-800-909-2  
Perfect score: 2468  
Sequence: 1 MAPVAVWAALAVGLELWAAA.....GSTEKPLPLGVDPAGMKPS 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2462	99.8	461	1	TNR2_HUMAN
2	1512	61.3	474	1	TNR2_MOUSE
3	379.5	15.4	349	1	VT22_VARV
4	360.5	14.6	326	1	VT2_MXXVL
5	341.5	13.8	435	1	TNR2_HUMAN
6	328	13.3	325	1	VT2_SFVKA
7	317.5	12.9	277	1	CD40_HUMAN
8	317	12.8	415	1	TNR2_MOUSE
9	259	10.5	283	1	TR14_HUMAN
10	252.5	10.2	289	1	CD40_MOUSE
11	251.5	10.2	416	1	NGFR_CHICK
12	242.5	9.8	269	1	CD40_BOVIN
13	241	9.8	595	1	CD30_HUMAN
14	229.5	9.3	271	1	OX40_RAT
15	223.5	9.1	272	1	OX40_MOUSE
16	217	8.8	277	1	OX40_HUMAN
17	210	8.5	427	1	NGFR_HUMAN
18	199	8.1	471	1	TNR1_BOVIN
19	198	8.0	461	1	TNR1_PIG
20	197	8.0	256	1	41BB_MOUSE
21	197	8.0	425	1	NGFR_RAT
22	191.5	7.8	454	1	TNR1_MOUSE
23	185.5	7.5	461	1	TNR1_RAT
24	185	7.5	255	1	41BB_HUMAN
25	183.5	7.4	455	1	TNR1_HUMAN
26	166.5	6.7	1367	1	AMYL_YEAST
27	152	6.2	327	1	FASA_BOVIN
28	148	6.0	327	1	FASA_MOUSE
29	147	6.0	600	1	SP96_DICDI
30	146.5	5.9	687	1	VS41_GIALA
31	144.5	5.9	835	1	CD97_HUMAN
32	144	5.8	332	1	FASA_PIG
33	143.5	5.8	324	1	FASA_RAT
					P20333 homo sapien
					P25119 mus musculus
					P34015 varliola vir
					P29825 myxoma viru
					P36941 homo sapien
					P25943 shope fibro
					P25942 homo sapien
					P50284 mus musculus
					Q92956 homo sapien
					P27512 mus musculus
					P18519 gallus gall
					Q28203 bos taurus
					P28908 homo sapien
					P15725 rattus norv
					P47741 mus musculus
					P43489 homo sapien
					P08138 homo sapien
					O19131 bos taurus
					P50555 sus scrofa
					P20334 mus musculus
					P07174 rattus norv
					P25118 mus musculus
					P22934 rattus norv
					Q07011 homo sapien
					P19438 homo sapien
					P08640 saccharomyc
					P51867 bos taurus
					P25446 mus musculus
					P14328 dictyosteli
					P92127 giardia lam
					P48960 homo sapien
					O77736 sus scrofa
					Q63199 rattus norv

34 142 5.8 723 1 DLLL\_HUMAN  
35 141.5 5.7 335 1 FASA\_HUMAN  
36 141.5 5.7 685 1 DLL4\_HUMAN  
37 141.5 5.7 1113 1 PER3\_MOUSE  
38 141 5.7 1255 1 MUC1\_HUMAN  
39 140.5 5.7 1959 1 AGRI\_RAT  
40 139.5 5.7 1169 1 YK82\_YEAST  
41 137.5 5.6 1696 1 PKK5\_BRACL  
42 135.5 5.5 775 1 ICP0\_HSV11  
43 135.5 5.5 796 1 YS8A\_CAEEL  
44 135.5 5.5 1955 1 AGRI\_CHICK  
45 135 5.5 677 1 T2D5\_HUMAN

## ALIGNMENTS

RESULT 1  
TNR2\_HUMAN  
ID TNR2\_HUMAN STANDARD; PRT; 461 AA.  
AC P20333;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR  
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).  
GN TNFRSF1B OR TNFR2 OR TNFR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90260639; PubMed=2160731;  
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,  
RA Dower S.K., Cosman D., Goodwin R.G.;  
RT "A receptor for tumor necrosis factor defines an unusual family of  
RT cellular and viral proteins.";  
RL Science 248:1019-1023(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91045991; PubMed=2172983;  
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,  
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;  
RT "A second tumor necrosis factor receptor gene product can shed a  
RT naturally occurring tumor necrosis factor inhibitor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96299745; PubMed=8661109;  
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,  
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,  
RA Brodeur G.M.;  
RT "Physical mapping and genomic structure of the human TNFR2 gene.";  
RL Genomics 35:94-100(1996).  
RN [4]  
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=90349572; PubMed=2166946;  
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,  
RA Ringold G.M.;  
RT "Complementary DNA cloning of a receptor for tumor necrosis factor  
RT and demonstration of a shed form of the receptor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).  
RN [5]  
RP SEQUENCE OF 27-31.  
RX MEDLINE=90110215; PubMed=2153136;  
RA Engelmann H., Novick D., Wallach D.;  
RT "Two tumor necrosis factor-binding proteins purified from human  
RT urine. Evidence for immunological cross-reactivity with cell surface  
RT tumor necrosis factor receptors.";  
RL J. Biol. Chem. 265:1531-1536(1990).  
RN [6]  
RP SEQUENCE OF 23-40; 65-69; 136-141; 300-306 AND 346-362.

RA MEDLINE=91056048; PubMed=2173696;  
RA Loetscher H., Schlaefer E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,  
RA Brockhaus M.;  
RT "Purification and partial amino acid sequence analysis of two  
RT distinct tumor necrosis factor receptors from HL60 cells.";  
RN J. Biol. Chem. 265:20131-20138(1990).  
RN [7]  
RP CHARACTERIZATION.  
RX MEDLINE=93016040; PubMed=1328224;  
RX Pennica D., Lam V.T., Mize N.K., Weber R.F., Fendly B.M.,  
RA Lipari M.T., Goeddel D.V.;  
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.  
RT Characterization of ligand binding, internalization, and receptor  
RT phosphorylation.";  
RL J. Biol. Chem. 267:21172-21178(1992).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.  
RX MEDLINE=99221490; PubMed=10206649;  
RX Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;  
RT "Structural basis for self-association and receptor recognition of  
RT human TRAF2";  
RL Nature 398:533-538(1999).  
CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND  
CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW  
CC LEVEL ON THREONINE RESIDUES.  
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND  
CC WYETH-AYERST). USED TO TREAT MODERATE TO SEVERE RHEUMATOID  
CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING  
CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO  
CC TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.  
CC -1- SIMILARITY: CONTAINS A LA-NGRP/TNFR-TYPE CYSTEINE-RICH REGION.  
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".  
CC -1- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;  
CC WWW="http://www.enbrelinfo.com/".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M32315; AAC59929.1; -;  
DR EMBL; M35857; AAC63262.1; -;  
DR EMBL; U52165; AAC50622.1; -;  
DR EMBL; U52156; AAC50622.1; JOINED.  
DR EMBL; U52157; AAC50622.1; JOINED.  
DR EMBL; U52158; AAC50622.1; JOINED.  
DR EMBL; U52159; AAC50622.1; JOINED.  
DR EMBL; U52160; AAC50622.1; JOINED.  
DR EMBL; U52161; AAC50622.1; JOINED.  
DR EMBL; U52162; AAC50622.1; JOINED.  
DR EMBL; U52163; AAC50622.1; JOINED.  
DR EMBL; U52164; AAC50622.1; JOINED.  
DR EMBL; M55994; AAC36755.1; -;  
DR PIR; A35356; A35356.  
DR PIR; A36007; A36007.  
DR PIR; A36475; A36475.  
DR PIR; B35010; B35010.  
DR PIR; A23666; A23666.  
DR PDB; 1CA9; 12-APR-99.  
DR MIM; 191191; -;  
DR InterPro; IPR001368; TNFR\_c6.  
DR Pfam; PF00020; TNFR\_c6; 4.  
DR ProDom; PD000771; TNFR\_c6; 1.  
DR SMART; SM00208; TNFR; 4.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
DR PROSITE; PS50050; TNFR\_NGFR\_2; 4.  
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;

KW Phosphorylation; Pharmaceutical; 3D-structure.  
FT SIGNAL 1 22  
FT CHAIN 23 461  
FT DOMAIN 23 257 TUMOR NECROSIS FACTOR RECEPTOR 2.  
FT TRANSMEM 258 287 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 288 461 POTENTIAL.  
FT DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 39 76 4 X TNFR-CYS.  
FT REPEAT 39 76 TNFR-CYS 1.  
FT REPEAT 77 118 TNFR-CYS 2.  
FT REPEAT 119 162 TNFR-CYS 3.  
FT REPEAT 163 201 TNFR-CYS 4.  
FT DISULFID 40 53 BY SIMILARITY.  
FT DISULFID 54 67 BY SIMILARITY.  
FT DISULFID 57 75 BY SIMILARITY.  
FT DISULFID 78 93 BY SIMILARITY.  
FT DISULFID 96 110 BY SIMILARITY.  
FT DISULFID 100 118 BY SIMILARITY.  
FT DISULFID 120 126 BY SIMILARITY.  
FT DISULFID 134 143 BY SIMILARITY.  
FT DISULFID 137 161 BY SIMILARITY.  
FT DISULFID 164 179 BY SIMILARITY.  
FT CARBOHYD 171 171 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 193 193 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 141 141 R -> P (IN REF. 4).  
FT CONFLICT 196 196 R -> M (IN REF. 1 AND 3).  
FT CONFLICT 363 363 A -> T (IN REF. 4).  
SQ SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;  
  
Query Match 99.8%; Score 2462; DB 1; Length 461;  
Best Local Similarity 99.8%; Pred. No. 3.9e-136;  
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MAPVAVWAALVGLLEWAAHAALPAQVATPYAPPEPGSTCLRLEYDQTAQMCCSKCSPG 60  
DB 1 MAPVAVWAALVGLLEWAAHAALPAQVATPYAPPEPGSTCLRLEYDQTAQMCCSKCSPG 60  
  
QY 61 QHAKVCTKTSVTDCSDSTYVTLWNWVPECLSCGSCSSDQVETQACTREQNRICTC 120  
DB 61 QHAKVCTKTSVTDCSDSTYVTLWNWVPECLSCGSCSSDQVETQACTREQNRICTC 120  
  
QY 121 RFGWCALSKQECRLCAPLRCRFGFVARPGTETSDVVKPCAPGTETSNSTSDICR 180  
DB 121 RFGWCALSKQECRLCAPLRCRFGFVARPGTETSDVVKPCAPGTETSNSTSDICR 180  
  
QY 181 PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPESTAPSTS 240  
DB 181 PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPESTAPSTS 240  
  
QY 241 FLLPMGPSPPAEGSTGDFALPVGLIVGVVGTALGLLIIGVNVCMVIMTOVKKPLCLQREKV 300  
DB 241 FLLPMGPSPPAEGSTGDFALPVGLIVGVVGTALGLLIIGVNVCMVIMTOVKKPLCLQREKV 300  
  
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSSALDRRAPTRNQPAQGVASGAGE 360  
DB 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSSALDRRAPTRNQPAQGVASGAGE 360  
  
QY 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSSSDHSSQCSSQASSTMGDTSSSPSESPEKDEQ 420  
DB 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSSSDHSSQCSSQASSTMGDTSSSPSESPEKDEQ 420  
  
QY 421 VPFSKECAFRSOLPETLLGSTEKPLPLGVDPAGMKPS 461  
DB 421 VPFSKECAFRSOLPETLLGSTEKPLPLGVDPAGMKPS 461  
  
RESULT 2  
TNR2\_MOUSE  
ID TNR2\_MOUSE STANDARD: PRT: 474 AA.  
AC P25119; P97893;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X69198; CAA49137.1; -;  
 DR EMBL; X67117; CAA47540.1; -;  
 DR PIR; D36858; D36858.  
 DR PIR; S35987; S35987.  
 DR PIR; S46888; S46888.  
 DR HSSP; P19438; INCF.  
 DR InterPro; IPR001368; TNFR\_c6.  
 DR Pfam; PF00020; TNFR\_c6; 2.  
 DR ProDom; PD000771; TNFR\_c6; 1.  
 DR SMART; SM00208; TNFR; 2.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00050; TNFR\_NGFR\_2; 2.  
 KW Repeat.  
 FT DOMAIN 31 108 2 X TNFR-CYS.  
 FT REPEAT 31 66 TNFR-CYS 1.  
 FT REPEAT 67 108 TNFR-CYS 2.  
 SQ SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;

Query Match 15.4%; Score 379.5; DB 1; Length 349;  
 Best Local Similarity 37.6%; Pred. No. 1.2e-15;  
 Matches 79; Conservative 24; Mismatches 86; Indels 21; Gaps 7;

QY 31 PYAPEGSTCLRYEYDQTAQMCKSCSPGQHAQVCTKTSVTCVDCSDSTVYQLNNV 90  
 DB 24 YPTP-PNGKCKDTEY--KRNLKLCCLSPGPGYASRLCDKNTQCTPCGSGTFTSRNHL 80  
 QY 91 PECLSCGSCSSDOVETQACTREONRICTCRPGWYCALSKQEGRLCAPLRCRPGFGVA 150  
 DB 81 PACLUSCGRNSQNSVETRSQNTNTHICECPGYCYLLKSGSGKACVSKQCGIGYGV 140  
 QY 151 RPTGTSDDVCKPCAPGTFSTNTSDICRPHQICNVVAIPGNA--SMDAVCT-----ST 203  
 DB 141 -GHTSVGVGVCSGCGFTGYSHTVSADKCEP-----VFNNTFNIDVETILYPVNDT 191  
 QY 204 SPTRSMAPGAVH--LPQPVSTRSQHTQPTP 231  
 DB 192 SCTRTTTTGLSESLTSELITMNHNTDCNP 221

RESULT 4  
 VT2\_MXXVL STANDARD; PRT; 326 AA.  
 AC P29825;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).  
 GN T2.  
 OS Myxoma virus (strain Lausanne).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Leporipoxvirus.  
 OX NCBI\_TaxID=31530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91335768; PubMed-1651597;  
 RA Upton C., Macen J.L., Schreiber M., McFadden G.;  
 RT "Myxoma virus expresses a secreted protein with homology to the tumor  
 RT necrosis factor receptor gene family that contributes to viral  
 RT virulence".  
 RL Virology 184:370-382(1991).  
 CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO  
 CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL  
 CC ANTIVIRAL EFFECTS OF THE CYTOKINE.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M95181; AAA46632.1; -;  
 DR EMBL; A23729; CAA01688.1; -;  
 DR PIR; A40566; GOV2ML.  
 DR HSSP; P19438; ITNR.  
 DR InterPro; IPR001368; TNFR\_c6.  
 DR Pfam; PF00020; TNFR\_c6; 2.  
 DR ProDom; PD000771; TNFR\_c6; 1.  
 DR SMART; SM00208; TNFR; 3.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00050; TNFR\_NGFR\_2; 2.  
 KW Receptor; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 326 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.  
 FT DOMAIN 27 186 4 X TNFR-CYS.  
 FT REPEAT 27 62 TNFR-CYS 1.  
 FT REPEAT 63 104 TNFR-CYS 2.  
 FT REPEAT 105 147 TNFR-CYS 3.  
 FT REPEAT 148 186 TNFR-CYS 4.  
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 14.6%; Score 360.5; DB 1; Length 326;  
 Best Local Similarity 36.9%; Pred. No. 1.4e-14;  
 Matches 75; Conservative 29; Mismatches 86; Indels 13; Gaps 6;  
 QY 31 PYAPEGSTCLRYEYDQTAQMCKSCSPGQHAQVCTKTSVTCVDCSDSTVYQLNNV 90  
 DB 20 PYGADRGR-CRGNDY--EKDGLCCTSCPGSVASRLCGPGSDTVCSCKNETFTASTNHA 76  
 QY 91 PECLSCGSCSSDOVETQACTREONRICTCRPGWYCALSKQEGRLCAPLRCRPGFGVA 150  
 DB 77 PACVSCRCRCCTGHLSESSQCDKTRDVCDSAGNYCLLKQEGCRICAPKCPAGYGV 136  
 QY 151 RPTGTSDDVCKPCAPGTFSTNTSDICRPHQICNVVAIPGN--ASMDAVCTSTSPTRS 208  
 DB 137 -GHTRTGDLVCTKPRYTSYDAVSTETCT--SSFNYISVEFNLYPVNDTSCTTTA---- 189  
 QY 209 MAPGAVHLPQPVSTRSQHTQPTP 231  
 DB 190 -GPNEVVKTSSEFVTLNHTDCDP 211

RESULT 5  
 TNRC\_HUMAN STANDARD; PRT; 435 AA.  
 ID TNRC\_HUMAN  
 AC P36941;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR  
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).  
 GN LTBR OR TNFR OR TNFRSF3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93252381; PubMed=8486360;  
 RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;  
 RT "Construction and evaluation of a hncDNA library of human 12p

RT transcribed sequences derived from a somatic cell hybrid.";

RN [2]

RP FUNCTION.

RX MEDLINE=94225209; PubMed=81711323;

RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,

RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;

RT "A lymphotoxin-beta-specific receptor.";

RL Science 264:707-710(1994).

CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN

CC IMMUNE DEVELOPMENT.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L04270; AAA36757.1; .

DR HSSP; P25942; 1CDF.

DR MIM; 600979; .

DR InterPro; IPR001368; TNFR\_c6.

DR Pfam; PF00020; TNFR\_c6; 4.

DR ProDom; PD000771; TNFR\_c6; 1.

DR SMART; SM00208; TNFR; 4.

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.

DR PROSITE; PS00505; TNFR\_NGFR\_2; 3.

KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.

FT SIGNAL 1 30

FT CHAIN 31 435

FT DOMAIN 31 227

FT TRANSMEM 228 248

FT DOMAIN 249 435

FT DOMAIN 42 211

FT REPEAT 42 81

FT REPEAT 82 124

FT REPEAT 125 168

FT REPEAT 169 211

FT DISULFID 43 58

FT DISULFID 59 72

FT DISULFID 62 80

FT DISULFID 83 98

FT DISULFID 101 116

FT DISULFID 104 124

FT DISULFID 126 132

FT DISULFID 139 148

FT DISULFID 142 167

FT DISULFID 170 185

FT CARBOHYD 40 40

FT CARBOHYD 177 177

FT SEQUENCE 435 AA; 46709 MW; 62462656022F656F CRC64;

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

FT N-LINKED (GLCNAC...) (POTENTIAL).

QY 167 GTFSTNTSSTDICRPHOICN---VVAIPGNASMDAVCTSTSPTRSMAPG-----AV 214  
Db 173 GHFQNTSSPSARCOPHTRCENOGVLEAPGTAQSDTTC--KNPLEPLPPENSGTMLLAV 230  
QY 215 HLPQP-----VST-----RSQHTQTPPEPSTAPSTFLL---PMGPSP-PAEGS----- 254  
Db 231 LLPLAPFLLLATVFCISWKS-----HPSLCRKGLSLLKRRPQEGPNPVAGSWEPPKA 283  
QY 255 -----TGDFALPVGLIIVGVTALGLLIIGVNVVIMTVQVKKKPLCLOREAKV 300  
Db 284 HPYFPDLVQPLLPISGDVS-PVS--TGLPAAPVLEAGVPQ-----QQSPDLDTRE--- 330  
QY 301 PHL-PADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRAPTRNOPOAPG 352  
Db 331 PQLPGEQSQVAHGTNGIHV-----TGGSMITIGNIYIYNGPVLGGPPGPG 376  
RESULT 6  
VT2\_SFVKA  
ID VT2\_SFVKA STANDARD; PRT; 325 AA.  
AC P25943;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).  
GN T2.  
OS Shope fibroma virus (strain Kasza) (SFV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Leporipoxvirus.  
OX NCBI\_TaxID=10272;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87321103; PubMed=2820128;  
RA Upton C., Delange A.M., McFadden G.;  
RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the  
RT telomeric region of the Shope fibroma virus genome.";  
RL Virology 160:20-30(1987).  
RN [2]  
RP FUNCTION.  
RX MEDLINE=91207415; PubMed=1850261;  
RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,  
RA McFadden G., Goodwin R.G.;  
RT "T2 open reading frame from the Shope fibroma virus encodes a soluble  
RT form of the TNF receptor.";  
RL Biochem. Biophys. Res. Commun. 176:335-342(1991).  
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO  
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL  
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.  
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query Match 13.8%; Score 341.5; DB 1; Length 435;  
Best Local Similarity 29.1%; Pred. No. 2.4e-13;  
Matches 120; Conservative 45; Mismatches 141; Indels 107; Gaps 25;  
QY 2 APVAVNAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56  
Db 9 APLAWGLVLGLFGLAASQPAVE-----PYASE-NQTCRQKEKYEYPOHRCSSR 61  
QY 57 CSPGQIAKVCFTKSTVDCSDSTSTYQLNWNVPCLSGSCSSDQV-----ETQACTR 112  
Db 62 CPPGTVVSACRSIRDTVCATCAENSYNHWNLYTICQLCR---PCDPVNGLEEIAFPTS 118  
QY 113 EQNRICTCRPGNYCALSKOEGCRCLAPKRCRPGFCVARGTSTSDVV-----CKPCAP 166  
Db 119 KRKTCRCOPGFMCAAWALE-CTHCELLSDCPG-----TEAEKDKVKGNNHCVPCKA 172





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:56:45 ; Search time 23.73 Seconds  
(without alignments)  
177.926 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_201\_257

Perfect score: 302

Sequence: 1 TSTSPTRSMAPGAVHLQPV.....STSFLLPMGPSPPAEGSTGD 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 288181

Minimum DB seq length: 0

Maximum DB seq length: 57

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	22.7	47	21 AAB56300	Human secreted pro
2	67.5	22.4	49	19 AAW59911	Amino acid sequenc
3	59	19.5	45	12 AAR13261	Linking B region #
4	58	19.2	54	15 AAR49722	Sequence of a pept
5	58	19.2	54	15 AAR49536	Camel Ig 2-heavy c
6	56.5	18.7	41	20 AAY01285	Peptide encoded by
7	56	18.5	47	22 AAM19336	Peptide #5770 enco
8	56	18.5	47	22 AAM32132	Peptide #6169 enco
9	55	18.2	46	20 AAW88522	Amphotropic hyperv
10	55	18.2	52	22 AAB61193	Human INTERCEPT 21
11	54.5	18.0	55	21 AAB42690	Human ORFX ORF2454

12	54	17.9	21	9 AAP81606	Sequence of human
13	54	17.9	36	21 AAY65289	Human 5' EST relat
14	54	17.9	50	22 AAM17997	Peptide #431 enco
15	54	17.9	50	22 AAM30506	Peptide #4543 enco
16	54	17.9	50	22 AAM05637	Peptide #4319 enco
17	53.5	17.7	37	18 AAW26641	H. insoliens family
18	52.5	17.4	21	14 AAR37992	Heavy chain hinge
19	52	17.2	50	18 AAW35739	Mucin peptide MUC1
20	52	17.2	51	18 AAW31697	Mucin peptide MUC1
21	51.5	17.1	55	22 AAM33030	Peptide #7067 enco
22	51	16.9	40	17 AAR88860	Peptide targeted
23	51	16.9	40	22 AAM18924	Peptide #5358 enco
24	51	16.9	40	22 AAM31492	Peptide #5529 enco
25	51	16.9	43	20 AAY49096	Amino acid sequenc
26	51	16.9	50	22 AAM20331	Peptide #6765 enco
27	51	16.9	50	22 AAM34632	Peptide #8669 enco
28	50	16.6	16	17 AAR99693	Iga protease speci
29	50	16.6	19	20 AAW99681	Human Iga1 hinge r
30	50	16.6	21	20 AAY42923	Core peptide in th
31	50	16.6	25	19 AAW64623	Human Iga1 CH2 reg
32	50	16.6	25	19 AAW61595	Human Iga1 CH2 rec
33	50	16.6	25	20 AAM85774	Iga1 CH2 region pr
34	50	16.6	52	21 AAG03497	Human secreted pro
35	50	16.6	52	22 AAG74844	Human colon cancer
36	49.5	16.4	40	16 AAR68002	Mucin repeat seque
37	49.5	16.4	41	19 AAW72703	Human mucin peptid
38	49.5	16.4	41	12 AAR15255	Linking B region #
39	49	16.2	27	9 AAP80541	Sequence of portio
40	49	16.2	35	15 AAR49650	Sequence of hinge
41	49	16.2	35	15 AAR49529	Camel 2-chain Ig h
42	49	16.2	40	19 AAW54873	Carcinoma-associat
43	49	16.2	46	20 AAY01284	Peptide encoded by
44	49	16.2	47	21 AAB38081	Fragment of human
45	48.5	16.1	34	20 AAY03809	Lck binding active

ALIGNMENTS

RESULT 1  
ID AAB56300 standard; Protein; 47 AA.  
XX AAB56300;  
XX AC  
XX 13-MAR-2001 (first entry)  
XX Human secreted protein sequence encoded by gene 99 SEQ ID NO:394.  
XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
KW antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;  
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virocid;  
KW fungicide; ophthalmological; gene therapy; pathological condition;  
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;  
KW cerebrovascular disorder; anglogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; skin aging; food additive; preservative.  
XX Homo sapiens.  
XX WC200070042-A1.  
XX 23-NOV-2000.  
XX 11-MAY-2000; 2000WO-US12788.  
XX Camel Ig 2-heavy c  
PR Peptide encoded by  
PR 13-MAY-1999; 99US-0134068.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;  
PI Duan RD, Florence KA, Soppet DR;



XX	OS	Camelus dromadarius.
XX	XX	
XX	FH	Key Location/Qualifiers
FT	Region	6..40 "hinge region"
FT	/note=	1..5
FT	Region	/note= "heavy chain variable region"
FT	/note=	41..54
FT	Region	/note= "heavy chain constant region"
XX	PN	EP584421-A.
XX	PD	02-MAR-1994.
XX	XX	
XX	Pf	21-AUG-1992; 92EP-04023326.
XX	PR	21-AUG-1992; 92EP-04023326.
XX	PA	(CAST/) CASTERMAN C.
XX	PA	(HAME/) HAMERS R.
XX	PI	Casterman C, Hamers R;
XX	XX	
XX	DR	WPI; 1994-067061/09.
XX	XX	
PT	PT	New isolated immunoglobulin molecules devoid of light polypeptide chains - consisting of heavy polypeptide chains only, obtd. from Camelid serum, for use as antibodies
PT	PT	
XX	XX	
PS	DS	Disclosure; Page 19; 35pp; English.
XX	XX	
CC	CC	This 2-chain Ig molecule lacks any light chains. The Ig has e.g. improved solubility and aggregate much less than heavy chains of 4-chain Igs. The Igs can be used normally for e.g. diagnosis, therapy, in vaccines, for isolation and purification of antigens and in the production of anti-idiotypic antibodies.
XX	SQ	Sequence 54 AA;
Query Match 19.2%; Score 58; DB 15; Length 54;		
Best Local Similarity 31.6%; Pred. No. 38;		
Matches 12; Conservative 3; Mismatches 23; Indels 0; Gaps 0;		
Qy	16	LPOPVYRSQHTPPPESTAPSTSFLPLMGSPSPAEG 53
	:	:           :
Dd	9	lpqpkpqppqpqqkpxpkpeptcpcpapellig 46
RESULT 6		
AAV01285		
ID	AAV01285	standard; Protein; 41 AA.
XX	AC	AAV01285;
XX	XX	
DT	01-JUN-1999	(first entry)
XX	XX	
DE	Peptide encoded by HRGP gene cassette.	
XX	XX	
KW	Synthetic gene; plant; gum; hydroxyproline-rich glycoprotein; HRGP;	
KW	repetitive proline-rich protein; RRP; arabinogalactan protein; AGP;	
KW	glycopeptide; internal repeat.	
XX	OS	Synthetic.
XX	PN	WO9903978-A1.
XX	PD	28-JAN-1999.
XX	PF	21-JUL-1998; 98WO-US15083.
XX	PR	20-JUL-1998; 98US-0897556.



AAB61193

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant; KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; KW hypotensive; dermatological; immunosuppressive; antiinflammatory; KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid; KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension; KW neurodegenerative disorder; osteoarthritis; graft vs host disease; KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; KW cholesterol ester storage; systemic lupus erythematosus; infection; KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma; KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; KW bone damage; cartilage damage; antiinflammatory disease; coagulation; KW thrombosis; contraceptive.

OS Homo sapiens.

XX W0200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI: 2000-602362/57.

DR N-PSDB; AAC76899.

XX Novel nucleic acids and peptides derived from open reading frame X, PT useful for treating e.g. cancers, proliferative disorders, PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 4083; 5507pp; English.

CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397, CC which represent the human ORFX open reading frames 1 to 3161. The ORFX CC sequences have activities such as: cytostatic; hepatotropic; vulnery; CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant; CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; CC antidiabetic; hypotensive; dermatological; immunosuppressive; CC antiinflammatory; antianaemic; antiviral; antifungal; antirheumatic; CC antithyroid; and antianaemic. The sequences can be used for determining CC the presence of or predisposition to, or preventing or treating CC pathological conditions associated with an ORFX-associated disorder. The CC nucleic acids can be used to express ORFX proteins in gene therapy CC vectors. The proteins and nucleic acids may be used to treat cancers, CC proliferative disorders, neurodegenerative disorders, osteoarthritis, CC graft vs host disease, cardiovascular disease, diabetes mellitus, CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, CC bacterial or fungal infection, malaria, autoimmune disorders, asthma, CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 55 AA;

Query Match 18.0%; Score 54.5; DB 21; Length 55;

Best Local Similarity 38.3%; Pred. No. 85;

Matches 18; Conservative 1; Mismatches 17; Indels 11; Gaps 3;

QY 19 PVSTRSQHTQTPERS-----TAPSTSFLLPMGPS---PPAEGSTG 56

Db 5 P\$g\$trtp\$tp\$pcphgdrllppsr--lpagpasafppaersrg 49

RESULT 12

AAP81606

ID AAP81606 standard; protein; 21 AA.

XX AAP81606;

XX 02-OCT-1990 (first entry)

DE Sequence of human immunoglobulin A1 (IgA1).

XX Human Cu/Zn superoxidisedismutase polymer; pSODCF1SODHA1;

KW human immunoglobulin A1; hinge region; antiinflammatory agent;

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 1..20

XX FT /note="Hinge #1"

XX FT Protein 4..20

XX FT /note="IgA1"

XX EP283244-A.

XX 21-SEP-1988.

XX 15-MAR-1988; 88EP-0302244.

XX 16-MAR-1987; 87US-0026143.

XX (CHIR-) CHIRON CORP.

XX Hallewell RA, Mullenbach G;

XX WPI: 1988-265657/38.

DR P-PSDB; AAN80435.

XX Superoxidisedismutase polymers having extended in vivo life - PT comprising superoxidisedismutase monomers covalently coupled by PT polypeptide spacers

XX Example; Page 5; 21pp; English.

XX A Cu/Zn superoxide dismutase (SOD) polymer of SOD units of at least 2 CC SOD monomers covalently coupled, carboxy terminus to amino terminus, to CC each other by a polypeptide spacer of at least 3 amino acids is claimed. CC The polymers are of formula

CC (SOD monomer - IgA - SOD monomer)<sub>x</sub> or

CC (SOD monomer - IgA - SOD monomer - IgA - SOD monomer)<sub>x</sub>

CC x = 2-4

CC IgA - a 10-100 amino acid long segment of an immunoglobulin hinge region CC AAN80435 was used to make pSODCF1SODHA1 encoding bacterial expression CC spacer-linked human SOD monomers. AAN80435 encodes human IgA1 hinge CC region beginning at AA residue 226 to avoid the Cys 225 residue with CC BamHI and NcoI sites at its ends. The SOD polymers have an extended CC circulatory life and retain the activity of SOD. The cloning and CC sequencing of human SOD (hSOD) cDNA and prodn. of hSOD in bacteria and CC yeast are described in EP-138111.

XX Sequence 21 AA;

Query Match 17.9%; Score 54; DB 9; Length 21;

Best Local Similarity 48.0%; Pred. No. 36;

Matches 12; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 16 LPQPVSTRSQHTQTPPEPSTAPSTS 40

Db 1 laqpvps----tpptpsptptps 21

RESULT 13





```
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PS analyzing gene expression in human placenta -
XX PS Claim 27; SEQ ID No 30775; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders.
XX SQ Sequence 50 AA;

Query Match 17.9%; Score 54; DB 22; Length 50;
Best Local Similarity 40.0%; Pred. No. 87;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 27 TQPTPEPSTAPSTSFLLPMGCPSPA 51
Db 17 tipleppasstppmlplpplppsps 41
```

Search completed: February 12, 2002, 12:59:58  
Job time: 193 sec

SQ SEQUENCE 415 AA; 44956 MW; 29B326A56AEF661 CRC64;  
  
 Query Match 12.8%; Score 317; DB 1; Length 415;  
 Best Local Similarity 24.9%; Pred. No. 6e-12;  
 Matches 119; Conservative 50; Mismatches 166; Indels 142; Gaps 22;  
  
 QY 7 WAALAVGLELWAAAHALPAQVAFPPYAPEGSTC--RLREYYDDTAQMCCSKSPGGHAK 64  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 14 WGPLLLGSLGLVA-SQPQLV--PPYRIE-NQTCDWDKKEYEPMDHVCCSGCPPEGFVF 69  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 65 VFCTKTSTVDSCDSEDTYQLNHWVECLISCGSGSSDQV----ETQACTREQRNRITCTC 120  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 70 AVCRSQDVTVCCKTCHPNYSNEHMNHLSCTCLR---PCDIVLGFEIVAPCTSDRKABCR 126  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 121 RPYGCALSQOEGRCRLCAPLRKCRPGFGVARPGTET-----SDVVCKPACPGETSNT 172  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 127 QPGMSCVYLNE-CVHCHEER-----LVLCOPGTAEAVTDIMTDCNVCPCKGFHFQNT 180  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 173 TSSDIDICRPHQICH-----VVAIPGNASMDAVCTSTSPTRSMAGAVHLPOPVSTRSOHTQ 228  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 181 SSPRARCOPTHCEIOGLVEAAPGTSYSDTICKN----- 214  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 229 PTPPSTAPSTSFLLPMGPSPBAEGSTGDPALVGLVGVLTALGLLIIGVNCVM---- 284  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 215 -PPPGAMLLAILLSL-----VLEFLTTVTLACAMMRHPS 249  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 285 -----TOVKK-----KPLCLQREAKVPHLP-----ADKARGTCQPEQHLLIT 322  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 250 LCRKLGLTKLRKHREGESPPCPPAPRAD-PHPDPDLAEPLPMSGDLSPAGPP-----T 302  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 323 APSSSSSLESSASALDRRAPTRNQOPAQGYEASGAGEARASTGSSSPGGHGCTQVNVT 382  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 303 APSLEEVLQQQSPLV-----QARELEAPEGEHQVAHCAN----GIHVTGGSVT 348  
 QY :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 383 CIVNVCSSDHSSOCSSQAESTMGDTOSSPESPKEQVPF--SKECAFRSQLETP 437  
 Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 349 VTGNIITYN-----GPVLGGT-RGPGDPPAPPPEPYTPEGAPGSELSTP 394  
  
 RESULT 9  
 TR14\_HUMAN STANDARD; PRT; 283 AA.  
 ID Q92956; Q9UM65;  
 AC 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 14 PRECURSOR  
 DE (HERPESVIRUS ENTRY MEDIATOR) (TUMOR NECROSIS FACTOR RECEPTOR-LIKE 2)  
 GN (TR2).  
 GE TNFRSF14 OR HVEM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RX [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Cervix adenocarcinoma;  
 RX MEDLINE=97033782; PubMed=8898196;  
 RA Montgomery R.I., Warner M.S., Lum B.J., Spear P.G.;  
 RT "Herpes simplex virus-1 entry into cells mediated by a novel member of  
 RT the TNF/NGF receptor family.";  
 RL Cell 87:427-436(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97306336; PubMed=9162061;  
 RA Kwon B.S., Tan K.B., Ni J., Oh K.-O., Lee Z.H., Kim K.K., Kim Y.-J.,  
 RA Wang S., Gentz R., Yu G.-L., Harrop J., Lyn S.D., Silverman C.,  
 RA Porter T.G., Truneh A., Young P.R.;  
 RT "A newly identified member of the tumor necrosis factor receptor  
 RT superfamily with a wide tissue distribution and involvement in  
 RT lymphocyte activation."  
 RL J. Biol. Chem. 272:14272-14276(1997).



RESULT 12



```

DR HSP; P19438; INCF.
DR MIM; 153243; -.
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00650; TNFR_NGFR_2; 2.
DR T-cell; Glycoprotein; Transmembrane; Signal; Receptor; Repeat;
KW Phosphorylation; Alternative Initiation.
FT SIGNAL 1 18
FT CHAIN 19 595
FT CHAIN 464 595
FT INIT MET 464 464
FT DOMAIN 19 379
FT TRANSMEM 380 407
FT DOMAIN 408 595
FT DOMAIN 28 325
FT REPEAT 28 66
FT REPEAT 68 106
FT REPEAT 107 150
FT REPEAT 205 241
FT REPEAT 243 281
FT REPEAT 282 325
FT DOMAIN 347 377
FT CARBOHYD 101 101
FT CARBOHYD 276 276
SQ SEQUENCE 595 AA; 63747 MW; 7A407CC78A6E0BC8 CRC64;

Query Match
Best Local Similarity 9.8%; Score 241; DB 1; Length 595;
Matches 113; Conservative 41; Mismatches 183; Indels 176; Gaps 17;

QY 11 AVGLELWAAHALPAQVATPYAPEGSTC--RLRYYDQTAQMCKSKSPGQHAKEFT 68
DB 7 ALGLFLGALRAFPQDRPE-----DICHGNPSHYDKAVRCCYRCPMGLFPTQQCP 59
QY 69 KTSDDTVCDSCEDSTYQLNWNVPECLUSGCRSSDDQVETQACTREONRICTRPGWYCAL 128
DB 60 QRPTDCRKOCEPDYLL---DEADRCTACVTCSSRDDLVKTPCAWNSRVCECRPGMFCST 116
QY 129 SKOEGRLCAPLRCRPGFGVARGPTGTSNVCKKCAPG----- 167
DB 117 SAVNSCARCFHSHVCPAGMIVKFPGTAQNTVCEPASPGVSPACASPENCKEPPSGTIPQ 176
QY 168 -----TFSNTT----- 173
DB 177 AKPTVPSPATSSASTNVPVGRGTRLAQEAASKLTRAPDSPSSVGRPSDDPGLSPTQCPGEG 236
QY 174 -----SSTDICRPHQICNVVAIPG 192
DB 237 SGDCRKOCEPDYLLDEAGRCTACVSCSRDDLVEKTPCAWNSRVCECRPGMFCST 296
QY 193 NASMDAVCTSTSTPSRMAGAVHLPO-----PVSTRSQHTQTPPE-----PST 235
DB 297 -----CARCPVPYPTCAAEVTKPDMAEKDTPFEAPPLGTQPD-CNPTPENGAEAPST 348
QY 236 APSTSFEL-----LPMGPSP-ASGSTGDFALPGLVIGVATGALLIIGVNVCMVTQ 286
DB 349 SPTQSLVDSQASKTLPIPTSAVALSSGKPVLDAGVFLFWILVLVVVGGSAFLCH 408
QY 287 VKKKPLCLQREAKVPHL--PADKARGTQGPQOHLITAPSSSSSSLESSASALDRAPT 344
DB 409 ---RRACRKRIQKHLCPVQV---TSQPKLE-LVDSRRPRSSSTQLRSGASVTEPVAAE 460
QY 345 R---NOPQAPGVASGAGERASTGSSDSPG 374
DB 461 RGLMSQPLMETCHSVGAAYLE-SLPLODASPAG 492

RESULT 14
OX40_RAT
ID OX40_RAT STANDARD; PRT; 271 AA.

```

```

AC P15725;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
GN TNFRSF4 OR TXGPIL OR OX40.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=T-cell;
RX MEDLINE=90214614; PubMed=2157591.
RA Mallett S., Fossum S., Barclay A.N.;
RT "Characterization of the MRC OX40 antigen of activated CD4 positive T
RT lymphocytes -- a molecule related to nerve growth factor receptor.";
RL EMBO J. 9:1063-1068(1990).
CC -|- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
CC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17037; CAA34897.1; -.
CC PIR; S08036; S08036.
CC PIR; S12783; S12783.
CC HSP; P25942; ICDF.
CC InterPro: IPR001368; TNFR_c6.
CC Pfam; PF00020; TNFR_c6; 3.
CC PRODOM; PD000771; TNFR_c6; 1.
CC SMART; SM00208; TNFR; 3.
CC PROSITE; PS00652; TNFR_NGFR_1; 3.
CC PROSITE; PS00650; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 271
FT DOMAIN 20 210
FT TRANSMEM 211 235
FT DOMAIN 236 271
FT DOMAIN 25 164
FT REPEAT 25 60
FT REPEAT 61 102
FT REPEAT 103 123
FT REPEAT 124 164
FT CARBOHYD 143 143
FT SEQUENCE 271 AA; 29895 MW; C06465136B16E821 CRC64;

Query Match
Best Local Similarity 9.3%; Score 229.5; DB 1; Length 271;
Matches 79; Conservative 28; Mismatches 88; Indels 97; Gaps 14;

QY 6 VW-----AALAVGLELWAAHALPAQVATPYAPEGSTRLREYD--QTAQMCKSKS 58
DB 3 VVWQPTAFLLGLSL-----GVTVKLNCVKDYTPSGHKCCREQ 42
QY 59 PGQHAKEFTKTSVDCSDSTYTQLNWN--VPECLUSGCRSSDDQVETQACTREQNR 116
DB 43 PGHGMVSRCDHTRDTVCHPCPEPGFYNEAVNYDTCKQCQNRSGSEL--KQCPTPTEDT 100
QY 117 ICTCRPGHYCALSKQEGCRCLCAPLRCRPGFGVARGPTGTS---DVVCKPCAPGTFESNT 173
DB 101 VCQCRPG-----TQPRDSSHKLGVDCVPCPGHFS---P 132

```

```
QY 174 SSTICRPHQICNV-----VAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT-- 227
D 133 GSNQACKPTWNTCTLGGKQIRHPASNSLDTVCEDRS-----LLATLWETQRTTF 181
QY 228 OPTPEPSTA-----PSTGELL-PMGSPSPAEGSTGDFALPVGLIVGVTA 270
D 182 RPTTVPTVTPWPTSQLPSTPTPLVAPEGPA-----FAVILGLGLLLA 224

RESULT 15
OX40_MOUSE
ID OX40_MOUSE STANDARD; PRT; 272 AA.
AC P47741;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TNFRSF4 OR TXGP1 OR OX40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=94044750; PubMed=8228223;
RA Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,
RA Claassen E., Noelle R.J., Fell H.;
RT "Cloning of mouse OX40: a T cell activation marker that may mediate
RT T-B cell interactions."
RL J. Immunol. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95255413; PubMed=7737295;
RA Birkeland M.B., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Barclay A.N.;
RT "Gene structure and chromosomal localization of the mouse homologue
RT of rat OX40 protein."
RL Eur. J. Immunol. 25:926-930(1995).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z21674; CAA79772.1; -
DR EMBL; X85214; CAA59476.1; -
DR HSP; P25942; ICDF.
DR MGD; MGI:104512; Tnfrsf4.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 272 OX40L RECEPTOR.
FT DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 212 236 POTENTIAL.
FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 165 4 X TNFR-CYS.
FT REPEAT 26 61 TNFR-CYS 1.
FT REPEAT 62 103 TNFR-CYS 2.
FT REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 125 165 TNFR-CYS 4.
```

```
FT CARBOHYD 144 144 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 15 15 A -> G (IN REF. 2).
SQ SEQUENCE 272 AA; 30153 MW; 06E7BB4156F0D08E CRC64;

Query Match 9.1%; Score 223.5; DB 1; Length 272;
Best Local Similarity 27.0%; Pred. No. le-06;
Matches 88; Conservative 32; Mismatches 97; Indels 109; Gaps 19;

QY 6 VW-----AALAVGLELWAAHALPAQVATPYAPEGSTCRLREYYDQTAQMCCSKCSPG 60
D 3 VWQOFTALLLLALTGLVTARRL-----NCVKHTY--PSGHKCCRECQPG 45
QY 61 QHAKVCTKTSYDSCDSTVTYQLWNW--VPECLSCGSRSCSDQVETQACTREQNRIC 118
D 46 HGMVSRCDHTRDTLCHPCETGYNEAVNYDTCKQCTQCNHRSGSEL--KQNCITPTQDTCV 103
QY 119 TCRPGWYCALSKQEGRCRLCAPLRKCRPGFVARPGTETS---DVCKPKCAPGTFSTNTSS 175
D 104 RCRPG-----TQPRQDSGVKLGVCVPCPPGHFS--PGN 135
QY 176 TDCRPHQICNV-----VAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQ-PT 230
D 136 NOACKPWTNCTLSGKQTRHPASDSLDVAVCED-----RSL-----LATLLWETQRP 181
QY 231 PEPSTA-----PSTSFLLPMGSPSPA-----EGSTGDFALPVGLIVGVTA-LGLLIIGVV 279
D 182 FRPTTVQSTTVWPRTSEL-----PSPPTLVTPGPA--FAVLGLGLGLLAPLVLL----- 231
QY 280 NCVIMTQVKKKPICLQREA-KVPHLP 304
D 232 -----ALYLLLRKAWRLPNT 246

Search completed: February 12, 2002, 12:51:31
Job time: 207 sec
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:47:13 ; Search time 28.69 Seconds  
(without alignments)  
2350.350 Million cell updates/sec

Title: US-09-800-909-2  
Perfect score: 2468  
Sequence: 1 MAPVAVWAALAVGLELWAAA.....GSTEEKPLPLGVPDAGMKRPS 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL17.\*
- 1: sp-archaea.\*
  - 2: sp-bacteria.\*
  - 3: sp-fungi.\*
  - 4: sp-human.\*
  - 5: sp-invertebrate.\*
  - 6: sp-mammal.\*
  - 7: sp-mhc.\*
  - 8: sp-organelle.\*
  - 9: sp-phage.\*
  - 10: sp-plant.\*
  - 11: sp-rodent.\*
  - 12: sp-virus.\*
  - 13: sp-vertebrate.\*
  - 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2279	92.3	425	4 Q16042	Q16042 homo sapien
2	1498	60.7	482	11 O88734	O88734 mus musculus
3	1477	59.8	459	11 O62327	O62327 mus musculus
4	826	33.5	161	4 Q9UJQ3	Q9UJQ3 homo sapien
5	503.5	20.4	175	11 Q9WUL4	Q9WUL4 rattus norv
6	482	19.5	92	4 Q9UIG9	Q9UIG9 homo sapien
7	409	16.6	78	4 Q9UIH0	Q9UIH0 homo sapien
8	390.5	15.8	349	12 O57101	O57101 monkeypox v
9	389.5	15.8	349	12 O57291	O57291 monkeypox v
10	389.5	15.8	349	12 O57102	O57102 monkeypox v
11	389	15.8	348	12 O57277	O57277 monkeypox v
12	389	15.8	348	12 O57103	O57103 monkeypox v
13	389	15.8	348	12 O57108	O57108 monkeypox v
14	386.5	15.7	349	12 O57099	O57099 monkeypox v
15	385	15.6	349	12 O57284	O57284 camelpox vi
16	385	15.6	349	12 O57098	O57098 camelpox vi
17	383.5	15.5	349	12 O57100	O57100 monkeypox v
18	381.5	15.5	349	12 O57111	O57111 variola vir
19	381.5	15.5	349	12 Q89098	Q89098 variola vir

20	381	15.4	347	12	O57115	O57115 cowpox viru
21	381	15.4	349	12	O57097	O57097 camelpox vi
22	379.5	15.4	348	12	O57112	O57112 variola vir
23	379.5	15.4	348	12	O85407	O85407 variola vir
24	378	15.3	347	12	O57119	O57119 cowpox viru
25	377	15.3	360	12	O57118	O57118 cowpox viru
26	376.5	15.3	349	12	O57110	O57110 variola vir
27	376.5	15.3	349	12	O89118	O89118 variola vir
28	374	15.2	351	12	O57117	O57117 cowpox viru
29	374	15.2	351	12	O57121	O57121 cowpox viru
30	373	15.1	351	12	O73559	O73559 cowpox viru
31	373	15.1	355	12	O85308	O85308 cowpox viru
32	371.5	15.1	349	12	O57109	O57109 variola vir
33	370.5	15.0	350	12	O57116	O57116 cowpox viru
34	368	14.9	349	12	O57305	O57305 cowpox viru
35	359	14.5	350	12	O57123	O57123 cowpox viru
36	354.5	14.4	326	12	O57122	O57122 cowpox viru
37	351.5	14.2	300	4	O95407	O95407 homo sapien
38	345.5	14.0	326	12	O57120	O57120 cowpox viru
39	343.5	13.9	655	4	O75509	O75509 homo sapien
40	339	13.7	316	12	O57092	O57092 ectromelia
41	339	13.7	320	12	O57091	O57091 ectromelia
42	339	13.7	320	12	O57300	O57300 ectromelia
43	330	13.4	320	12	O57079	O57079 cowpox viru
44	330	13.4	322	12	O72761	O72761 cowpox viru
45	330	13.4	372	4	Q9UHP4	Q9UHP4 homo sapien

ALIGNMENTS

RESULT 1

Q16042	ID	Q16042	PRELIMINARY;	PRT;	425 AA.
AC	Q16042;				
DT	01-NOV-1996	(TREMBlrel. 01, Created)			
DT	01-JUN-2001	(TREMBlrel. 17, Last sequence update)			
DT	01-JUN-2001	(TREMBlrel. 17, Last annotation update)			
DE	TUMOR NECROSIS FACTOR RECEPTOR (FRAGMENT).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91370690; PubMed=1966549;				
RA	Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,				
RA	Brockhaus M., Lesslauer W.;				
RT	"Two human TNF receptors have similar extracellular, but distinct				
RL	intracellular, domain sequences.";				
RT	Cytokine 2:231-237(1990).				
DR	EMBL;	S63368;	AA019824.2;	-;	
DR	HSSP;	P25942;	ICDF		
DR	InterPro;	IPR001368;	TNFR_c6.		
DR	Pfam;	PF00020;	TNFR_c6;	4.	
DR	ProDom;	PD000771;	TNFR_c6;	1.	
DR	SMART;	SM00208;	TNFR;	4.	
DR	PROSITE;	PS00652;	TNFR_NGFR_1;	1.	
DR	PROSITE;	PS50050;	TNFR_NGFR_2;	3.	
KW	Receptor.				
FT	NON_TER	1			
SQ	SEQUENCE	425 AA;	44608 MW;	1B24A97E3AD4CF9F	CRC64;

Query Match 92.3%; Score 2279; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 4.4e-168;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 GSTCRLREYYDQTQACMCCSKCSPGOHAKVFCFKTSDTVCDSCEDSTYTQLWNVPCLSC 96  
Db 1 GSTCRLREYYDQTQACMCCSKCSPGOHAKVFCFKTSDTVCDSCEDSTYTQLWNVPCLSC 60  
Qy 97 GSRCSDDVQACTREQNRICTRCPGWYCALSKQEGRLCAPLRKCRPGFGVARGTGT 156



```
Db 61 CADCEASMTQVWVNFRTCLSCSSCSTDOVETRACTKQNRVYCAEAGRYCALKTHSGS 120
QY 134 CRICAPLRCRPGFVGVARGTETSDVVCKPCAPGTSTNTSDICRPHQICNVVAIPGN 193
Db 121 CROCMLRSCGPGFVGASSRAPNGVNLCKACAPGTSDTSTSDVCRPHRICSLAIPGN 180
QY 194 ASMDAVCTSTSPTRSMAGVAVHLPQPVSTRSOHTQTPPEPSTAPSTFLLPMGPPSAPG 253
Db 181 ASDAVCAPESPPLSAIPRTLVYSQPEPTRSQPLDQEPGSPFP--SILTSLGSTPIEQ 238
QY 254 ST-GDPALPVGLIVGTALGLLIIVVNCVIMTVQVKKPLCLQREAKVPHLPADKARGTQ 312
Db 239 STYKGLISLPIGLIVGTSGLLMLVNCFLIVQRKKPSCLORDAKVPHVPEKSDAV 298
QY 313 GPEQQHLLTAPSSSSSSLESSASALDRAPTRNQPAQGV--EASGAGEARASTGSSDSS 371
Db 299 GLEQQHLLTAPSSSSSSLESSASAGDRAPPQGGHGPQARYMAEAQGSQEARASSRISDSS 358
QY 372 PGHGTQVNVTCIVNVCSDDHSSQSSQASSTMGTDTSPPSPKDEQVFPFSKECAFR 431
Db 359 HGSCHTHVNVTCIVNVCSDDHSSQSSQASATVGDPAKPSAPKDEQVFPFSQEECPQ 418
QY 432 SQLETPETLLGSTEEKPLPLGVDPDAGMKPS 461
Db 419 SPYETETEL--QSHKPLPLGVDPDAGMKPS 446

RESULT 4
Q9UJ03 ID Q9UJ03 PRELIMINARY; PRT; 161 AA.
AC Q9UJ03
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE DJ111824.3 (TUMOR NECROSIS FACTOR RECEPTOR 2 (75 KDA) (TNF BINDING
PROTEIN 2, TBPII, TNF-R2, CD120B, TNFR)) (FRAGMENT).
GN TNFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Heath P.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031276; CAA20343.1; -.
FT NON_TER 1
SQ SEQUENCE 161 AA; 16449 MW; E6C96E774B5CF7B0 CRC64;

Query Match 33.5%; Score 826; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.8e-56;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSLESSASALDRAPTRNQPAQGVASGAGE 360
Db 1 PHLPADKARGTQGPQEQHLLITAPSSSSLESSASALDRAPTRNQPAQGVASGAGE 60
QY 361 ARASTGSSDSSPGHGTQVNVTCIVNVCSDDHSSQSSQASSTMGTDTSPPSPKDEQ 420
Db 61 ARASTGSSDSSPGHGTQVNVTCIVNVCSDDHSSQSSQASSTMGTDTSPPSPKDEQ 120
QY 421 VPFSKECAFRSQLETPETLLGSTEEKPLPLGVDPDAGMKPS 461
Db 121 VPFSKECAFRSQLETPETLLGSTEEKPLPLGVDPDAGMKPS 161

RESULT 5
Q9WUL4 ID Q9WUL4 PRELIMINARY; PRT; 175 AA.
AC Q9WUL4
DT 01-NOV-1999 (TReMBLrel. 12, Created)
```

```
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR TYPE II (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=OVARY;
RA Balchak S.K., Marcinkiewicz J.L.;
RT "Evidence for the presence of Tumor Necrosis Factor Alpha Receptors
During Ovarian Development in the Rat.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF142499; AAD30148.1; -.
DR InterPro; IPR001368; TNFR_C6.
DR SMART; SM00208; TNFR; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 175
SQ SEQUENCE 175 AA; 18201 MW; 499EADAAB21ED8B CRC64;

Query Match 20.4%; Score 503.5; DB 11; Length 175;
Best Local Similarity 55.9%; Pred. No. 1.4e-31;
Matches 99; Conservative 23; Mismatches 52; Indels 3; Gaps 2;

QY 147 FGVARPGTSDVVCKPCAPGTSTNTSDICRPHQICNVVAIPGNASMDAVCTSTSP 206
Db 1 FGVASSTSGNVCISACAPGTSDTSTSDVCRPHRICSLAIPGNASTDAVCASEPT 60
QY 207 RSMAPGAVHLPQPVSTRSQHTQTPPEPSTAPSTFLLPMGPPSPPAEGS-TGDFALPVGLI 265
Db 61 PSAGPTIYVSQPEPTRSQPMGQEPGSPQTSHIP--VSLGSTPIIEPSITGGISLPIGLI 118
QY 266 VGV TALGLLIIGVNVCMITQVKKPKLCQREAKVPHLPADKARGTQGPQEQHLLIT 322
Db 119 VGLTTLGLMLGLATCFILVQRKKPKSCLQRETMVPHLPDDKSDQAIGLEQGHLLTT 175

RESULT 6
Q9UIG9 ID Q9UIG9 PRELIMINARY; PRT; 92 AA.
AC Q9UIG9
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 (FRAGMENT).
GN TNFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Komata T., Tsuchiya N., Matsushita M., Tokunaga K.;
RT "New polymorphism within the extracellular region of TNFR2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030952; BAA89055.1; -.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 92 AA; 9530 MW; 89BDE40B7CC4FE1 CRC64;

Query Match 19.5%; Score 482; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 3.5e-30;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 SSPGCHGTQVNVTCIVNVCSDDHSSQSSQASSTMGTDTSPPSPKDEQVFPFSKECA 429
Db 1 SSPGCHGTQVNVTCIVNVCSDDHSSQSSQASSTMGTDTSPPSPKDEQVFPFSKECA 60
QY 430 FRSQLETPETLLGSTEEKPLPLGVDPDAGMKPS 461
```

```

Db      61  FRSQLETPTLLGSTEEKPLPLGVDPDAGMKPS 92
|||||
RESULT 7
Q9UIH0 PRELIMINARY; PRT; 78 AA.
ID Q9UIH0
AC Q9UIH0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 (FRAGMENT).
GN TNFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Komata T., Tsuchiya N., Matsushita M., Tokunaga K.;
RT "New polymorphism within the extracellular region of TNFR2.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB030951; BAA89054.1; -
KW Receptor.
FT NON_TER 1
FT VARIANT 48
FT NON_TER 78
FT NON_TER 78
SQ SEQUENCE 78 AA; 7841 MW; 3A219A37EAFE0719 CRC64;

Query Match 16.6%; Score 409; DB 4; Length 78;
Best Local Similarity 98.7%; Pred. No. 1.3e-24;
Matches 77; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 185 CNVVAIPGNASMDAVCTSTSTRSMAPCAVHLPOPVSTRSHTOPTPEPSTAPSTSLLP 244
|||||
DB 1 CNVVAIPGNASMDAVCTSTSTRSMAPCAVHLPOPVSTRSHTOPTPEPSTAPSTSLLP 60
|||||
QY 245 MGPSPPAEGSTGDFALPV 262
|||||
DB 61 MGPSPPAEGSTGDFALPV 78
|||||

RESULT 8
O57101 PRELIMINARY; PRT; 349 AA.
ID O57101
AC O57101;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ZAIRE-1977 (77-0666);
RL Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U87845; AAB94362.1; -
DR HSSP: P25942; 1CDF.
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 2.
DR ProDom: PD000771; TNFR_c6; 1.
DR SMART: SM00208; TNFR; 2.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
SQ SEQUENCE 349 AA; 38311 MW; 02F65B00CFB858BE CRC64;

Query Match 15.8%; Score 390.5; DB 12; Length 349;

```

```

Best Local Similarity 41.8%; Pred. No. 1.4e-22;
Matches 77; Conservative 21; Mismatches 71; Indels 15; Gaps 5;

Qy 31 PYAPEPGSTCLRLRYDQTAQMCCSKSPGOHAKVFTKTSDTVCDSCEDSTYTQLWNV 90
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 24 PHAPNGK-CKDNEY--RSRLNCLCLSCPPGTIASRLCDSKNTQCTPCGSDTFTSHNHL 80

Qy 91 PECLSCGRSSDQVETQACTREONRICTCRPGWYCALSKQEGCRLCAPLKRCPGFGVA 150
   |||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 81 QACLSCNGRCDNSQVETRSCNTTHNRICESPGYYCLLKGSNGCRTCSIKTKGIGYGV 140

Qy 151 RPGTETSDVVKCPKCAPGTFSNTTSSDICRPHQICN-----VVAIPGNASMDAVCTS 202
   |||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 141 -GYTSTGDVICSPPGPGTYSHTVSSTDKCEPVVTSNTFNVIDEINLYPVN---DTSCTR 196

Qy 203 TSPT 206
   |::|
Db 197 TTTT 200

RESULT 9
O57291 PRELIMINARY; PRT; 349 AA.
AC O57291;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TUNOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RC SEQUENCE FROM N.A.
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88144; AAB94369.1; -.
DR EMBL; U87842; AAB94359.1; -.
DR EMBL; U87994; AAB94365.1; -.
DR EMBL; U87995; AAB94366.1; -.
DR EMBL; U88143; AAB94368.1; -.
DR HSP; P25942; LCDF.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 2.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00208; TNFR; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
SQ SEQUENCE 349 AA; 39295 MW; CBD2C949ED2B8E7C CRC64;

Query Match 15.8%; Score 389.5; DB 12; Length 349;
Best Local Similarity 41.8%; Pred. No. 1.7e-22;
Matches 77; Conservative 21; Mismatches 71; Indels 15; Gaps 5;

Qy 31 PYAPEPGSTCLRLRYDQTAQMCCSKSPGOHAKVFTKTSDTVCDSCEDSTYTQLWNV 90
   |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 24 PHAPNGK-CKDNEY--RSRLNCLCLSCPPGTIASRLCDSKNTQCTPCGSDTFTSHNHL 80

Qy 91 PECLSCGRSSDQVETQACTREONRICTCRPGWYCALSKQEGCRLCAPLKRCPGFGVA 150
   |||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 81 QACLSCNGRCDNSQVETRSCNTTHNRICESPGYYCLLKASGCRCTCSIKTKGIGYGV 140

Qy 151 RPGTETSDVVKCPKCAPGTFSNTTSSDICRPHQICN-----VVAIPGNASMDAVCTS 202
   |||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 141 -GYTSTGDVICSPPGPGTYSHTVSSTDKCEPVVTSNTFNVIDEINLYPVN---DTSCTR 196

Qy 203 TSPT 206
   |::|
Db 197 TTTT 200

```

Query Match<sup>c</sup> 15.8%; Score 390.5; DB 12; Length 349;

```
DR Pfam; PF00020; TNFR_c6; 2.  
DR ProDom; PD000771; TNFR_c6; 1.  
DR SMART; SM00208; TNFR; 2.  
DR PROSITE; PS00652; TNFR_NGFR_1; 2.  
DR PROSITE; PS50050; TNFR_NGFR_2; 2.  
SQ SEQUENCE   348 AA;  33212 MW;  54019521556C2D8F CRC64;  
  
Query Match           15.8%; Score 389; DB 12; Length 348;  
Best Local Similarity 42.1%; Pred. No. 1.8e-22;  
Matches    77; Conservative      22; Mismatches     70; Indels    14; Gaps          6;  
  
Qy 31 PYAEPGSTCLRLRYDYDTQMCCSKSPGOHAKVFTKTDVTDCSCEDSITYTQLNNWV 90  
Db :||| ||| |:::|| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
24 PHAPSNKG-KKDNEY--RSRNLCLSLCPGGPGTYASRLCDSDKNTQTCTPCGSDFTFSHNHL 80  
Yy :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
91 PECLSCGRSSDOVEQTACTREONRICTCPGWYCALSKEOECGLCAPLRKCPRGFVA 150  
Db :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
81 QACLSNGCRSDNQVFTRSCNTTHNRICEGPSGYICLLKGSSGCRTCLISKTKCGIGCVS 140  
Qy 151 RPYGTETSDVVCKPACPAGTSNTSTSDICTRP--HQICNVIAI-----PGNASMDAVCTST 203  
Db :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
141 -GYTSTGDVICSPCGPGTYSHTVSVSTDKEPEVTSNTFNIDVELINLPVN---DTSCIRT 196  
CrMB.  
Qy 204 SPT 206  
Db :|  
197 TTT 199  
  
RESULT 12  
OS07103 PRELIMINARY; PRT; 348 AA.  
AC OS7103;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DE DE  
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.  
GN CRMB.  
QS Monkeypox virus.  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxviriinae;  
OC Orthopoxvirus.  
NCBI_Taxid=10244;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=ZAIRE-1979 (79-0005);  
RA Loparev V.N., Parsons J.M.; Esposito J.J.;  
KL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.  
RR EMBL; U87847; AAB94364.1; -.  
DR HSSP; P25942; ICDF.  
DR InterPro; IPR001368; TNFR_c6.  
DR Pfam; PF00020; TNFR_c6; 2.  
Dr ProDom; PD000771; TNFR_c6; 1.  
DR SMART; SM00208; TNFR; 2.  
DR PROSITE; PS00652; TNFR_NGFR_1; 2.  
DR PROSITE; PS50050; TNFR_NGFR_2; 2.  
SQ SEQUENCE   348 AA;  38184 MW;  34A5E6B8B27907B5 CRC64;  
  
Query Match           15.8%; Score 389; DB 12; Length 348;  
Best Local Similarity 42.1%; Pred. No. 1.8e-22;  
Matches    77; Conservative      22; Mismatches     70; Indels    14; Gaps          6;  
  
Qy 31 PYAEPGSTCLRLRYDYDTQMCCSKSPGOHAKVFTKTDVTDCSCEDSITYTQLNNWV 90  
Db :||| ||| |:::|| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
24 PHAPSNKG-KKDNEY--RSRNLCLSLCPGGPGTYASRLCDSDKNTQTCTPCGSDFTFSHNHL 80  
Yy :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
91 PECLSCGRSSDOVEQTACTREONRICTCPGWYCALSKEOECGLCAPLRKCPRGFVA 150  
Db :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
81 QACLSNGCRSDNQVFTRSCNTTHNRICEGPSGYICLLKGSSGCRTCLISKTKCGIGCVS 140  
Qy 151 RPYGTETSDVVCKPACPAGTSNTSTSDICTRP--HQICNVIAI-----PGNASMDAVCTST 203  
Db :||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||  
141 -GYTSTGDVICSPCGPGTYSHTVSVSTDKEPEVTSNTFNIDVELINLPVN---DTSCIRT 196
```

[illegible]

Db 80 I<sup>1</sup>PACLS<sup>1</sup>CNG<sup>1</sup>GRD<sup>1</sup>SNQ<sup>1</sup>YET<sup>1</sup>RSC<sup>1</sup>NT<sup>1</sup>TH<sup>1</sup>RICE<sup>1</sup>CPG<sup>1</sup>YIC<sup>1</sup>LK<sup>1</sup>SSG<sup>1</sup>CKAC<sup>1</sup>VSOT<sup>1</sup>KCGIG<sup>1</sup>GV<sup>1</sup> 139.  
 Qy 150 ARPGT<sup>1</sup>ETSD<sup>1</sup>VVK<sup>1</sup>KPC<sup>1</sup>AGT<sup>1</sup>FSNT<sup>1</sup>STD<sup>1</sup>CRHQ<sup>1</sup>ICNV<sup>1</sup>VAIP<sup>1</sup>GN<sup>1</sup>A-----SMD 197  
 Db 140 S-GHT<sup>1</sup>SAGD<sup>1</sup>ICSP<sup>1</sup>CGLT<sup>1</sup>YSRT<sup>1</sup>SSAD<sup>1</sup>KCEP-----VPSNT<sup>1</sup>FNYID<sup>1</sup>VEIN<sup>1</sup>LYPVND 190  
 Qy 198 AVCTST<sup>1</sup>SPT 206  
 Db 191 TSC<sup>1</sup>TRTT<sup>1</sup>TT 199

Search completed: February 12, 2002, 12:50:47  
Job time: 214 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 12, 2002, 12:44:58 ; Search time 24.82 Seconds  
(without alignments)  
1375.817 Million cell updates/sec

Title: US-09-800-909-2  
Perfect score: 2468  
Sequence: 1 MAPVAVWAALAVGLELWAAA.....GSTEEKPLPLGVDPAGMKPS 461

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2468	100.0	461	AA111141	Human TNF-R deduce
2	2468	100.0	461	AA111141	Fibroblast derived
3	2468	100.0	461	AA111141	A human tumour nec
4	2462	99.8	461	AA111001	40KD TNF inhibitor
5	2462	99.8	461	AA111001	Human tumour necro
6	2462	99.8	461	AA111001	Death receptor. H
7	2462	99.8	461	AA111001	Human TNF receptor
8	2462	99.8	461	AA111001	Human tumour necro
9	2462	99.8	461	AA111001	Human 40 kDa TNF i
10	2394	97.0	461	AA111001	p75 Tumour Necrosi
11	2376	96.3	461	AA111001	Sequence of human

12	2051.5	83.1	392	20	AA130935	Human tumour necro
13	2050.5	83.1	392	12	AA11605	Human 75KD TNF-bin
14	1506	61.0	474	12	AA111142	TNF-R deduced from
15	1433	58.1	518	22	AA170001	STNFR(075):Fc fusi
16	1433	58.1	518	22	AA150080	TNFR:Fc fusion pro
17	1424	57.7	518	15	AA151003	Sequence of a reco
18	1404	56.9	485	13	AA124016	Fusion protein TNF
19	1381	56.0	248	21	AA194718	Human type 2 tumou
20	1318	53.4	235	21	AA154443	Wild type N-termina
21	1315	53.3	235	21	AA154440	Amino acid sequenc
22	1315	53.3	235	21	AA154441	Amino acid sequenc
23	1312	53.2	235	19	AA159665	Human soluble tumo
24	1312	53.2	235	19	AA152270	Tumour necrosis fa
25	1312	53.2	235	20	AA189234	Tumour necrosis in
26	1312	53.2	235	21	AA154442	A K108R/K120R muta
27	1312	53.2	235	22	AA137685	Human 40 kDa TNF i
28	1263	51.2	227	22	AA166981	Tnfr2 protein. Un
29	1256	50.9	225	21	AA177463	Primate protein se
30	1116	45.2	198	21	AA194720	Human type 2 tumou
31	1028	41.7	183	16	AA177421	BamTP delta53 nerv
32	937	38.0	165	21	AA180014	Peptide fragment o
33	935	37.9	163	21	AA194712	Tumour necrosis fa
34	856	34.7	159	22	AA137683	Human 40 kDa TNF i
35	794.5	32.2	258	22	AA150082	Rat TNFR (p80) ext
36	794.5	32.2	487	22	AA150084	TNFR:Fc fusion pro
37	770	31.2	227	21	AA177462	Truncated protein seq
38	707	28.6	122	19	AA152271	Truncated soluble
39	695	28.2	120	22	AA166991	Human TNF-II pepti
40	551	22.3	93	22	AA137684	Human 40 kDa TNF i
41	477	19.3	802	16	AA170111	TBPII-GBP 130 fusi
42	470	19.0	802	16	AA170112	TNF-R-GBP 130 fusi
43	453	18.4	77	20	AA194641	TNF-R extracellular
44	453	18.4	77	22	AA169193	Human TNF-R extrac
45	379.5	15.4	349	22	AA150523	Human tumour necro

## ALIGNMENTS

RESULT 1  
AA111141  
ID AA111141 standard; Protein; 461 AA.  
AC AA111141;  
XX  
XX 24-MAY-1991 (first entry)  
XX  
XX Human TNF-R deduced from clone 1.  
XX  
XX Tumour necrosis factor receptor; immune response; inflammation;  
KW cachexia; septic shock.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..22  
FT /label= signal sequence  
FT Protein 23..461  
FT /label= TNF receptor  
FT Domain 258..287  
FT /label= transmembrane region  
XX  
XX EP418014-A.  
XX  
XX 20-MAR-1991.  
XX  
XX 10-SEP-1990; 90EP-0309875.  
XX  
XX 10-MAY-1990; 90US-0523635.  
XX  
XX 11-SEP-1989; 89US-0405370.  
XX  
XX 13-OCT-1989; 89US-0421417.  
XX  
XX (IMMU-) IMMUNEX CORP.



```

XX Smith CA, Goodwin RG, Beckmann PM;
PI WPI: 1991-082230/12.
DR N-PSDB; AAQ10990.
XX
XX New tumour necrosis factor -alpha and -beta receptors - and DNA
PT encoding these used to regulate immune responses in treatment of
PT cachexia, septic shock or side-effects of cytokine therapy.
XX
XX Disclosure; Fig 2; 4lpp; English.
XX
XX The sequence was deduced from a DNA sequence obtd. from a clone
CC isolated from a library prepd. from a human fibroblast cell line,
CC WT-26 V44 (ATCC CCL 95.1). The clone is deposited as Accession No.
CC 68088 under the name pCAV/NOT-TNF-R. The DNA can be truncated to
CC produce sequences which express soluble receptor comprising
CC residues 1-235, 1-185 or 1-163 of the protein.
XX See also AAR11142.
XX
XX Sequence 461 AA;
SQ

```

Query Match 100.0%; Score 2468; DB 12; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 1e-152;  
 Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFPTYPAPGPCSTCLREYYDQTQMCCSKSPG 60
DB 1 mapvavwaalavglelwaahaalpaqvafptypapGPCSTCLREYYDQTQMCCSKSPG 60
QY 61 QHAKVCTKTSDFVCDSCEDSTVTLNWNVPECLSCGSRSSDOVETQACTREONR1CTC 120
DB 61 qhakvctktsdfvcdscedstyqlwnvpeclscgssrdovetqactreognrictc 120
QY 121 RPYWYCALSKQEGRCRLCAPLKRCPGFGVAPGTETSDVVKCPACPGTFSNTTSSTDICR 180
DB 121 rpywycalskqegrcrlcaplkrpGFGVAPGTETSDVVKCPACPGTFSNTTSDICR 180
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240
DB 181 phqicnvvaipgnasmdavctstsptrsmagavhlpqpvstrsqhtqptpepstapsts 240
QY 241 FLPLPMGPSPAEGSTGDFALPVGLVGVTAIGLLIIGVNCVIMTVQKKPLCLQREAKV 300
DB 241 flplpmgpSPAEGSTGDFALPVGLVGVTAIGLLIIGVNCVIMTVQKKPLCLQREAKV 300
QY 301 PHLPAKARQTQGPQOHLITAPSSSSSSLESSASALDRRAPTRNOPAGVEASGAGE 360
DB 301 phlpakarqtqGPQOHLITAPSSSSSSLESSASALDRRAPTRNOPAGVEASGAGE 360
QY 361 ARASTGSSDSSPGHGTVQNVNVCIVNVCSSDHSSQCSASSTMGDTDSPSPSPKDEQ 420
DB 361 arastgssdSPHGTVQNVNVCIVNVCSSDHSSQCSASSTMGDTDSPSPSPKDEQ 420
QY 421 VPESKECAPRSOLETPETLLGSTEELPLGLVPDAGMKPS 461
DB 421 vpeskecaprsOLETPETLLGSTEELPLGLVPDAGMKPS 461

```

RESULT 2  
 AAR42058  
 ID AAR42058 standard; Protein; 461 AA.  
 XX  
 AC AAR42058;  
 XX  
 DT 29-APR-1994 (first entry)  
 XX  
 DE Fibroblast derived TNF-R.  
 XX  
 KW Human; tumour necrosis factor receptor; TNF-R; interleukin-1 receptor;  
 KW IL-1R; fusion protein; linker; TNF; IL-1; cachexia; cerebral malaria;  
 KW rheumatoid arthritis; diabetes; multiple sclerosis; septic shock;

```

KW pulmonary fibrosis; silicosis; allograft; xenograft; rejection;
KW graft verses host disease; sepsis; inflammation; allergy;
KW autoimmune dysfunction.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..22 /note= "Signal peptide"
FT Protein 23..461 /note= "Mature hTNF-R"
FT Peptide 23..185 /note= "Preferred soluble TNF-R"
FT Peptide 23..207 /note= "Preferred soluble TNF-R"
FT Peptide 23..257 /note= "Preferred soluble TNF-R"
FT Peptide 23..206 /note= "Preferred soluble TNF-R"
FT Peptide 23..204 /note= "Preferred soluble TNF-R"
FT Peptide 1..206 /note= "Preferred soluble TNF-R"
FT Peptide 1..204 /note= "Preferred soluble TNF-R"
XX WO9319777-A.
XX 14-OCT-1993.
XX 26-MAR-1993; 93WO-US02938.
XX 30-MAR-1992; 92US-0860710.
XX (IMMUNEX CORP.
XX Smith CA;
XX WPI: 1993-336592/42.
DR N-PSDB; AAQ49931.
XX
XX New fusion protein tumour necrosis factor and human interleukin-1
PT receptor - useful in therapy, diagnosis and assays of e.g.
PT rheumatoid arthritis, diabetes, cerebral malaria, sepsis, etc.
XX
XX Claim 5; Fig 2; 85pp; English.
XX
XX The sequences given in AAR42058-59 represent human tumour necrosis
CC factor receptor (TNF-R) and the sequences in AAR42060-61 represent
CC human interleukin-1 receptor (IL-1R). These sequences were used in
CC the production of a fusion protein which conformed to one of the
CC formulae:
CC TNF-R-linker-TNF-R-linker-IL-1R
CC IL-1R-linker-TNF-R-linker-TNF-R or
CC TNF-R-linker-TNF-R
CC
CC The linker may comprise 5-100 amino acids selected from Gly, Asp,
CC Ser, Thr and Ala. These linkers separate the individual moieties
CC by such a distance that each component of the fusion protein is
CC capable of folding into the secondary or tertiary structure required
CC for its biological activity. These fusion proteins may be used in
CC therapy, diagnosis and assays for conditions mediated by TNF or IL-1,
CC particularly in conditions in which both TNF and IL-1 play a causative
CC role. They may be used to treat cachexia, rheumatoid arthritis,
CC diabetes, multiple sclerosis, pulmonary fibrosis and silicosis,
CC cerebral malaria, allograft and xenograft rejection in graft verses
CC host disease, sepsis, septic shock, inflammation, allergies and
CC autoimmune dysfunctions.
XX
XX Sequence 461 AA;
SQ

```

Query Match 100.0%; Score 2468; DB 14; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 1e-152;

```
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPVAVMAALAVGLELWAAHAHALPAQVATPYAPPEGSTCRLEYYDQTAQMCCKSPG 60
DB 1 mapvavmaalavglelwaahaalpaqvafpyapegstcrleyydqtaqmcckspg 60
QY 61 QHAKVFCTKTSQVCDSCDSTYTLQNMWVPECLSGSRCSDDQVETQACTRQNRITC 120
DB 61 qhakvfctktsdvcscdstytlqnmwvpeclsgsrscsddvetqactreqnrietc 120
QY 121 RFGWYCALSKQEGCRCLCAPLRKCRPGFVGVARPGTETSDVWCKPCAPGTFSTNTSTDICR 180
DB 121 rfgwycalskqegcrclcaplrkcrpgfvgvarpgtetstdvwcpcapgtfstntstdicr 180
QY 181 PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPQPVSTRSQHTOPTPEPSTAPSTS 240
DB 181 phqicnvvaipgnasmdavctstsptrsmapgavhlpqpvstrsqhtoptpepstapsts 240
QY 241 FLPLMGPSPPAEGSTGDFALPVLGVLIVGTALGLLIIGVNVCMVMTQVKKKPLCLQREAKV 300
DB 241 flplmgpsppaegstgdfalpvlgvliivgtalgliliigvncvmvmtqvkkkplclqreakv 300
QY 301 PHLPADKARGTQGEQHLITAPSSSSSSLESSASALDRRAPTRNQPAQPGVEASGAGE 360
DB 301 phlpadkargtggeqhlitapsssssslessasaldraptrnqpapqpgveasgag 360
QY 361 ARASTGSSDSSPGCGHGTQVNVTCIVNVCSDDHSSQCSQASMTGDTSSSPSEPKDEQ 420
DB 361 arastgssdspgghgtqvnvtcivnvcsddhssqcsqasmtgdtsspsespkd 420
QY 421 VPFSKEECAPRSQLETPETLLGSTEEKPLPLGVDPDAGMKPS 461
DB 421 vpfskeecafrsqletpetllgsteeekplplgvdpdagmkps 461

RESULT 3
AAB18717
ID AAB18717 standard; Protein: 461 AA.
XX
AC AAB18717;
XX
DT 22-JAN-2001 (first entry)
XX
DE A human tumour necrosis factor family receptor (TNF-RII).
XX
KW Human; tumour necrosis factor family receptor; TRI; tumour growth;
KW cell proliferation; chlamydia infection; immunodeficiency; septic shock;
KW T-cell mediated autoimmune disease; acquired immunodeficiency syndrome;
KW AIDS; cerebral malaria; graft rejection; cytotoxicity; cachexia;
KW apoptosis; inflammation; cancer; cardiovascular disease; angiogenesis;
KW inflammatory disease; atherosclerosis; diabetes mellitus; allergy;
KW neurological disorder; autoimmune disease; wound healing; bone formation;
KW osteoporosis.
XX
OS Homo sapiens.
XX
PN WO200054651-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-006592.
XX
PR 15-MAR-1999; 99US-0124489.
PR 26-MAY-1999; 99US-0136248.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Greene JM, Fleischmann RD, Ni J;
XX
DR WPI; 2000-618858/59.
XX
PT Novel tumour necrosis factor family receptor for diagnosing and treating
PT acquired immunodeficiency syndrome, cancer, cardiovascular diseases,
```

```
PT Inflammatory diseases and autoimmune diseases -
XX
PS Claim 16; Page 24-25; 228pp; English.
XX
CC The present sequence encodes human tumour necrosis factor family
CC receptor (TNF-RII) polypeptide. The specification describes a TNF
CC receptor designated TRI. An agonist to the TRI receptor is useful for
CC inhibiting tumour growth, to stimulate human cellular proliferation, to
CC regulate immune response and antiviral response, to protect against the
CC effects of ionising radiations, to protect against chlamydia infections,
CC to regulate growth, and to treat immunodeficiencies such as in human
CC immunodeficiency virus (HIV). An antagonist to the TRI receptor is
CC useful for treating T-cell mediated autoimmune diseases, acquired
CC immunodeficiency syndrome (AIDS), septic shock, cerebral malaria, graft
CC rejection, cytotoxicity, cachexia, apoptosis, and inflammation. TRI
CC polynucleotides and polypeptides, and TRI agonists and antagonists
CC are useful for treating cancers, cardiovascular diseases, inflammatory
CC diseases, atherosclerosis, diabetes mellitus, neurological disorders,
CC autoimmune diseases, for promoting angiogenesis, for treating allergy,
CC for wound healing, for regulating bone formation and for treating
CC osteoporosis.
XX
SQ Sequence 461 AA;
Query Match 100.0%; Score 2468; DB 21; Length 461;
Best Local Similarity 100.0%; Pred. No. 1e-152;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPVAVMAALAVGLELWAAHAHALPAQVATPYAPPEGSTCRLEYYDQTAQMCCKSPG 60
DB 1 mapvavmaalavglelwaahaalpaqvafpyapegstcrleyydqtaqmcckspg 60
QY 61 QHAKVFCTKTSQVCDSCDSTYTLQNMWVPECLSGSRCSDDQVETQACTRQNRITC 120
DB 61 qhakvfctktsdvcscdstytlqnmwvpeclsgsrscsddvetqactreqnrietc 120
QY 121 RFGWYCALSKQEGCRCLCAPLRKCRPGFVGVARPGTETSDVWCKPCAPGTFSTNTSTDICR 180
DB 121 rfgwycalskqegcrclcaplrkcrpgfvgvarpgtetstdvwcpcapgtfstntstdicr 180
QY 181 PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPQPVSTRSQHTOPTPEPSTAPSTS 240
DB 181 phqicnvvaipgnasmdavctstsptrsmapgavhlpqpvstrsqhtoptpepstapsts 240
QY 241 FLPLMGPSPPAEGSTGDFALPVLGVLIVGTALGLLIIGVNVCMVMTQVKKKPLCLQREAKV 300
DB 241 flplmgpsppaegstgdfalpvlgvliivgtalgliliigvncvmvmtqvkkkplclqreakv 300
QY 301 PHLPADKARGTQGEQHLITAPSSSSSSLESSASALDRRAPTRNQPAQPGVEASGAGE 360
DB 301 phlpadkargtggeqhlitapsssssslessasaldraptrnqpapqpgveasgag 360
QY 361 ARASTGSSDSSPGCGHGTQVNVTCIVNVCSDDHSSQCSQASMTGDTSSSPSEPKDEQ 420
DB 361 arastgssdspgghgtqvnvtcivnvcsddhssqcsqasmtgdtsspsespkd 420
QY 421 VPFSKEECAPRSQLETPETLLGSTEEKPLPLGVDPDAGMKPS 461
DB 421 vpfskeecafrsqletpetllgsteeekplplgvdpdagmkps 461

RESULT 4
AAB11001
ID AAB11001 standard; Protein: 461 AA.
XX
AC AAB11001;
XX
DT 13-MAY-1991 (first entry)
XX
DE 40kD TNF inhibitor precursor.
XX
KW Tumour necrosis factor; Inhibitor.
```

XX OS Homo sapiens.  
XX PN AU9058976-A.  
XX PD 24-JAN-1991.  
XX PF 16-JUL-1990; 90AU-0058976.  
XX PR 07-FEB-1990; 90US-0479661.  
XX PR 18-JUL-1989; 89US-0381080.  
XX PR 11-DEC-1989; 89US-0450329.  
XX PA (SYNE-) SYNERGEN INC.  
XX WPI; 1991-073847/11.  
XX DR N-PSDB; AAQ10907.  
XX PT Tumour necrosis factor inhibitor - for suppression of TNF-alpha  
XX and -beta, useful as therapeutic agent.  
XX PS Disclosure; Fig 39; 142pp; English.  
XX CC The sequence comprises the entire 40 kD TNF inhibitor. The clone  
CC from which the sequence was deduced was isolated from a cDNA  
CC library prep. from RNA form U937 cells treated with PMA/PHA.  
CC The whole gene can be inserted into expression vectors for prepn.  
CC of TNF inhibitor for use in the treatment of inflammatory and  
CC degenerative diseases.  
CC See also AARI0986 and AARI0984.  
XX SQ Sequence 461 AA;

Query Match 99.8%; Score 2462; DB 12; Length 461;  
Best Local Similarity 99.8%; Pred. No. 2.6e-152;  
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVAAALAVGLELWAAAHALPAQVAFTPYAPGPGSTCLRREYDQTAQMCCSKCSPG 60  
DB 1 mapvavaaalavglelwaaahalpavaftpyapepgstclrrreydydqtaqmccskcspg 60  
QY 61 QHAKVFTKTSDRVCDSCEDSTYTQLMNNWVPECLSCGRSSDQVETQACTREQNRICTC 120  
DB 61 qhakvftktsdrtvcdscdstytqlwnvpeclscgrsscdqvetqactreqnrtctc 120  
QY 121 RPYGWCALSKQEGRCRLCAPLRKCRPGFGVAPGTETSDVVKCPACPTFSNTTSDICR 180  
DB 121 rpygwcalskqegrcrlcaplrkcrpgfgvapggtetdvvkcpacptfsnttsdclcr 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQVPVSTRSOHTQPTPEPSTAPSTS 240  
DB 181 phqicnvvaipgnasmdavctstsptrsmapgavhlpqvpvstrsqhctqptpepstapsts 240  
QY 241 FLPLPMGFPSPAEGSTGDFALFVLGVGTALGLLIIGVWNCVIMTVQKKPLCLQREAKV 300  
DB 241 flplpmgfpspaegstgdfalpvglvgvtalglilgvncvintvqkkplclqreakv 300  
QY 301 PHLPADKARTQGPQEQHLLITAPSSSSSLESSASALDRAPTRNQPQAPGVEASGAGE 360  
DB 301 phlpadkargtgpqeqhllitapsssslessasaldraptrnqpqpgveasgag 360  
QY 361 ARASTGSSDSSPGHGTCVQNVNVCVSSDHSSQSSQASSTMGDTSDSPSPKDEQ 420  
DB 361 arastgssdsspghgctvqnvntcivncssdhssqssqassmgdtsdsspspkdeq 420  
QY 421 VPFSKECAPRSOLETPETLLGSTEERKPLPLGVDPDAKMKPS 461  
DB 421 vpfskecafrsletpetllgsteerkplplgvdpdagmkps 461

RESULT,  
AAB37801

ID AAB37801 standard; Protein; 461 AA.  
XX AC AAB37801;  
XX XX 23-FEB-2001 (first entry)  
XX DE Human tumour necrosis factor p75 receptor.  
XX KW Human; tumour necrosis factor; TNF; TNF-alpha; TNF-beta; p55 receptor;  
KW p75 receptor; antiinflammatory; haemostatic; antibacterial; sepsis;  
KW immunosuppressive; immunomodulator; cardiant; cytostatic; cachexia;  
KW neuroprotective; respiratory; inflammation; infection; Crohn's disease;  
KW multiple sclerosis; autoimmune disorder; cardiovascular disorder;  
KW chronic myelogenous leukaemia; inflammatory bowel disease.  
XX OS Homo sapiens.  
XX PN WO2000064479-A1.  
XX PD 02-NOV-2000.  
XX PF 26-APR-2000; 2000WO-US11700.  
XX PR 27-APR-1999; 99US-0301274.  
XX PA (ANTI-) ANTIBODY SYSTEMS INC.  
XX PI Fredeking TM, Ignatyev GM;  
XX WPI; 2000-679646/66.  
XX Novel compositions comprising tetracycline or tetracycline-like  
XX compounds for the treatment and/or prevention of acute inflammatory  
XX responses and diseases, e.g. septic shock and immune complex-induced  
XX colitis -  
XX Disclosure; Page 169-171; 183pp; English.  
XX CC The present sequence is given in a specification relating to novel  
XX compositions and methods containing tetracycline or tetracycline-like  
XX compounds for treating and/or preventing acute inflammatory responses and  
XX diseases. Such diseases include acute inflammatory conditions associated  
XX with viral haemorrhagic diseases (including diseases caused by  
XX Bunyaviridae, Filoviridae, Flaviviridae or Arenaviridae viruses),  
XX parasitic diseases, bacterial infections, sepsis, cachexia, autoimmune  
XX disorders, acute cardiovascular events, chronic myelogenous leukaemia and  
XX transplanted bone marrow-induced graft-versus-host disease, septic shock,  
XX immune complex-induced colitis, cerebrospinal fluid inflammation,  
XX multiple sclerosis, inflammatory responses associated with trauma,  
XX systemic inflammatory response syndrome (SIRS), adult respiratory  
XX distress syndrome (ARDS), acute liver failure, inflammatory bowel disease  
XX and Crohn's disease.  
XX SQ Sequence 461 AA;

Query Match 99.8%; Score 2462; DB 21; Length 461;  
Best Local Similarity 99.8%; Pred. No. 2.6e-152;  
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVAAALAVGLELWAAAHALPAQVAFTPYAPGPGSTCLRREYDQTAQMCCSKCSPG 60  
DB 1 mapvavaaalavglelwaaahalpavaftpyapepgstclrrreydydqtaqmccskcspg 60  
QY 61 QHAKVFTKTSDRVCDSCEDSTYTQLMNNWVPECLSCGRSSDQVETQACTREQNRICTC 120  
DB 61 qhakvftktsdrtvcdscdstytqlwnvpeclscgrsscdqvetqactreqnrtctc 120  
QY 121 RPYGWCALSKQEGRCRLCAPLRKCRPGFGVAPGTETSDVVKCPACPTFSNTTSDICR 180  
DB 121 rpygwcalskqegrcrlcaplrkcrpgfgvapggtetdvvkcpacptfsnttsdclcr 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPSTS 240

```

Db 181 phqicnvvaipgnasrdavctstsptrmapgavhlpqpvstrsqtqtpepstapsts 240
QY 241 FLPLPKCPSPAPGSGTGFALPVLGVVTAALGLLIIGVNCVIMTQVKKKPLCLQREAKV 300
Db 241 flpmpgspapagsgtdfaldpvlvgvtafgllilggnvncvmtqgkkkplclqreakv 300
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSLESSASALDRRAPTRNOAPGVEASGAGE 360
Db 301 phlpadkargtgqpeqghllitapsssslessasaldrraptrnqpapqveasgag 360
QY 361 ARASTGSSDSPGCHGTQVNVTCIVNVCSSSHSSQCSQASSTMGDTSSPSEPKDEQ 420
Db 361 arastgssdspgchgtqvnvtcivnvcssdhssqcsqasstmgdtsspsespkd 420
QY 421 VPFSKECAFRSQLETPETLLGSTEKPLPLGVPDAGMKPS 461
Db 421 vpfskecafrsqletpetllgsteekplplgvpdagmkps 461

```

## RESULT 6

AAB01342  
ID AAB01342 standard; Protein: 461 AA.

AC AAB01342;

DT 25-SEP-2000 (first entry)

DE Death receptor.

KW UL144; death receptor; apoptosis; programmed cell death; FAS;  
KW TNF-R1; TRAMP; DR-6; TRAIL; modulation; treatment; cancer; virus;  
KW human.

OS Homo sapiens.

PN WO200034335-A2.

PD 15-JUN-2000.

PF 03-DEC-1999; 99WO-US26035.

PR 04-DEC-1998; 98US-0205018.

PA (SCHE ) SCHERING CORP.

PI Leong C, Phillips JH;

DR WPI; 2000-423383/36.

PT Purified or recombinant polypeptide for modulating apoptosis comprises  
PT a sequence which binds to an antibody specific for UL144 or its  
PT fragments

PS Disclosure; Page 74-75; 76pp; English.

CC A pure or recombinant polypeptide which binds to a polyclonal antibody  
CC specific for the mature UL144 is useful for screening molecules which  
CC block induction of apoptosis or interfere with antiapoptotic activity.  
CC The polypeptide is also useful for modulating apoptosis and useful in  
CC treatment of conditions associated with abnormal physiology or  
CC development, such as cancer or degenerative conditions and for  
CC regulation of viral infection and replication. At least five  
CC different death receptors are known, which include the CD95  
CC (Fas/APO-1), the TNF receptor-1, TNF receptor apoptosis-mediated  
CC protein (TRAMP), death receptor-6 (DR-6), and TNF-related  
CC apoptosis-inducing ligand (TRAIL) receptors 1, 2 and 4.

SQ Sequence 461 AA;

Query Match 99.8%; Score 2462; DB 21; Length 461;  
Best Local Similarity 99.8%; Pred. No. 2.6e-152;

```

Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAPVAVAAALAVGLELWAAAHALPAQVAFTPYAPEPGSTGRLREYDQTAOMCCSKCSPG 60
Db 1 mapvavaaalavglelwaahalpaqvaftpyapepgstgrrlreydqtacmccskcspg 60
QY 61 QHAKVFTKTSDTVCDSCEDSTYTQLMNWVPECLSGSRCSDDQVETQACFREQNRITC 120
Db 61 qhakvftktsdtvcdscdstytqlmnwvpeciscgrcsddqvetaqreqrictc 120
QY 121 RPYWYCALSKQEGRLCAPLRKCRPGFVARPGTETSDVCKPCAPGTFNTTSTDTICR 180
Db 121 rpywycalskqegrlcaplrkcrpgfvarpgtetsdvckpcapgtfnttsatdicer 180
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQTPPEPSTAPSTS 240
Db 181 phqicnvvaipgnasrdavctstsptrsmapgavhlpqpvstrsqhtqtppepstapsts 240
QY 241 FLPLPMGPSPAPGSGTGFALPVLGVVTAALGLLIIGVNCVIMTQVKKKPLCLQREAKV 300
Db 241 flpmpgspapagsgtdfaldpvlvgvtafgllilggnvncvmtqgkkkplclqreakv 300
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSLESSASALDRRAPTRNOAPGVEASGAGE 360
Db 301 phlpadkargtgqpeqghllitapsssslessasaldrraptrnqpapqveasgag 360
QY 361 ARASTGSSDSPGCHGTQVNVTCIVNVCSSSHSSQCSQASSTMGDTSSPSEPKDEQ 420
Db 361 arastgssdspgchgtqvnvtcivnvcssdhssqcsqasstmgdtsspsespkd 420
QY 421 VPFSKECAFRSQLETPETLLGSTEKPLPLGVPDAGMKPS 461
Db 421 vpfskecafrsqletpetllgsteekplplgvpdagmkps 461

```

## RESULT 7

AAB35331  
ID AAB35331 standard; Protein: 461 AA.

AC AAB35331;

DT 08-MAY-2001 (first entry)

DE Human TNF receptor SEQ ID NO: 6.

KW Human; tumour necrosis factor receptor; TR13; TR14; infection;  
KW cancer; autoimmune disease; allergy; inflammatory disease;  
KW graft rejection; apoptosis; cardiovascular disease; aneurysm.

OS Homo sapiens.

PN WO200105834-A1.

PD 25-JAN-2001.

PF 14-JUL-2000; 2000WO-US19343.

PR 16-JUL-1999; 99US-01444087.

PR 18-AUG-1999; 99US-0149450.

PR 20-SEP-1999; 99US-0149712.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Young PE;

DR WPI; 2001-112682/12.

PT Nucleic acids encoding 2 human tumor necrosis factor receptor  
PT polypeptides (TR13) and (TR14), useful for the prevention, diagnosis  
PT and treatment of, e.g. cancers, acquired immune deficiency syndrome and  
PT hypohidrotic ectodermal dysplasia -  
XX

PS Disclosure; Page 377-378; 418pp; English.

XX The present invention provides the protein and coding sequences of the

CC human tumour necrosis factor receptors TR13 and TR14. These sequences are

CC useful in the diagnosis and treatment of many diseases, including cancer,

CC autoimmune diseases, cardiovascular disorders, allergies, aneurysms and

CC neurodegenerative diseases, graft rejection, inflammation, cancer;

XX infections.

XX Sequence 461 AA;

QY Query Match 99.8%; Score 2462; DB 22; Length 461;

DB Best Local Similarity 99.8%; Pred. No. 2.6e-152;

XX Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVAAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLEYYDQTAQMCCSKCSPG 60

DB 1 mapvavaaalavglelwaahahalpaqvaftpyapepgstcrleyydqtacmccskcspg 60

QY 61 QHAKVFCFKTSDTVCDCSDSTYTLWNWVPECLSCGSRCSDDVETOACTREONRICTC 120

DB 61 qhakvfcfktsdvtcdscdstytlwnwvpeclscgsrcsddvetoactregnrictc 120

QY 121 RPYWCALSKOEGRCILCAPLRCRPGFVARPGTETSDVWCKPCAPGTFTSNTSTDICR 180

DB 121 rpywcalskqegrcilcaplrcrpgfvarpgtetsdvwckpcapgtftsntstdicr 180

QY 181 PHQICNVVAIPGNASMDAVCTSTPTRSMAPGAVHLPQVSTRSQHTQTPPEPSTAPSTS 240

DB 181 phqicnvvaipgnasmdavctstptrsmapgavhlppqvstrsqhtqtppepstapsts 240

QY 241 FLPMGSPPAEGSTGDFALPVLGVVGTALGLLIIGVNCVIMTVQVKKPLCLQREAKV 300

DB 241 flpmgspspaegstgdfalpvlgvvgtalgliliigvncvimtvqvkklclqreakv 300

QY 301 PHLPADKARGTQGPQOQHLLITAPSSSSLESSASALDRRAPTRNQPAQGVASGAGE 360

DB 301 phlpadkargtgqgpqoqhllitapsssslessasaldraptrnqpqagveasgag 360

QY 361 ARASTGSSDSSPGHGHTQVNVTCIVNVCSSDHSQSSQASSTMGDTDSSPSPKDEQ 420

DB 361 arastgssdsspgghgtqvnvtcivnvcssdhsqssqasstmgdtdsspspkdeq 420

QY 421 VPSKECAPRSOLETETLLGSTEKPLPLGVDPAGMKPS 461

DB 421 vpskecafrsletetllgstekplplgvdpagmkps 461

RESULT 8

AAB36698

ID AAB36698 standard; Protein: 461 AA.

XX AC

XX AAAB36698;

XX 15-MAR-2001 (first entry)

XX Human tumour necrosis factor receptor TNFR2 protein SEQ ID NO:3.

XX Human; tumour necrosis factor receptor 5; TR1D; TNFR-5; TR5; neurotropic;

KW TRAIL receptor without intracellular domain; diagnosis; cytostatic;

KW tumour necrosis factor related apoptosis inducing ligand; vasotropic;

KW immunosuppressive; neuroprotective; antiviral; antiinflammatory;

KW anticonvulsant; antiparasitic; cardiant; anti-HIV; antiparkinsonian;

KW gene therapy; restenosis; graft versus host disease; tumour; cancer;

KW apoptotic cell death related disease; autoimmune disorder;

KW cardiovascular disorder; viral infection.

XX Homo.sapiens.

OS

XX WO200071150-A1.

PN

XX 30-NOV-2000.

PD

XX 18-MAY-2000; 2000WO-US13515.

XX 20-MAY-1999; 99US-0135164.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Wei Y, Ruben SM, Gentz RL, Ni J;

XX WPI; 2001-041051/05.

DR Nucleic acid encoding a TR1D polypeptide, also referred to as tumor

XX necrosis factor receptor 5, useful in the diagnosis, treatment or

PT prevention of cancer, autoimmune disorders and viral infection -

PT Disclosure; Fig 2; 285pp; English.

XX The present invention describes the human TR1D protein (tumour necrosis

CC factor (TNF) related apoptosis inducing ligand (TRAIL) receptor without

CC intracellular domain, also referred to as tumour necrosis factor

CC receptor 5 (TNFR-5 or TR5)). TR1D has cytostatic, immunosuppressive,

CC neurotropic, neuroprotective, antiviral, antiinflammatory, anticonvulsant,

CC antiparasitic, cardiant, anti-HIV, antiparkinsonian and vasotropic

CC activities, and can be used in gene therapy. The TR1D polynucleotides

CC are useful for detecting complementary polynucleotides. TR1D proteins and

CC polynucleotides are useful in the treatment of tumours, resistance to

CC parasite, bacteria and viruses, restenosis and graft versus host disease.

CC They are also useful for inducing proliferation of T-cells, endothelial

CC cells and certain haematopoietic cells, to regulate antiviral responses

CC and to prevent certain autoimmune diseases after stimulation of TR1D by

CC an agonist or TRAIL binding facilitator. The antibodies which bind TR1D

CC polypeptides are useful for treating and/or preventing diseases

CC associated with increased or decreased apoptotic cell death. The TR1D

CC polynucleotides, proteins, antibodies, agonists and antagonists are

CC useful in the diagnosis, treatment or prevention of: (a) cancer;

CC (b) autoimmune disorders; (c) diseases associated with increased

CC apoptosis; (d) cardiovascular disorders; and (e) viral infection. The

CC present sequence represents a tumour necrosis factor receptor used in

CC comparison with TR1D in the exemplification of the present invention.

XX Sequence 461 AA;

QY Query Match 99.8%; Score 2462; DB 22; Length 461;

DB Best Local Similarity 99.8%; Pred. No. 2.6e-152;

XX Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVAAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLEYYDQTAQMCCSKCSPG 60

DB 1 mapvavaaalavglelwaahahalpaqvaftpyapepgstcrleyydqtacmccskcspg 60

QY 61 QHAKVFCFKTSDTVCDCSDSTYTLWNWVPECLSCGSRCSDDVETOACTREONRICTC 120

DB 61 qhakvfcfktsdvtcdscdstytlwnwvpeclscgsrcsddvetoactregnrictc 120

QY 121 RPYWCALSKOEGRCILCAPLRCRPGFVARPGTETSDVWCKPCAPGTFTSNTSTDICR 180

DB 121 rpywcalskqegrcilcaplrcrpgfvarpgtetsdvwckpcapgtftsntstdicr 180

QY 181 PHQICNVVAIPGNASMDAVCTSTPTRSMAPGAVHLPQVSTRSQHTQTPPEPSTAPSTS 240

DB 181 phqicnvvaipgnasmdavctstptrsmapgavhlppqvstrsqhtqtppepstapsts 240

QY 241 FLPMGSPPAEGSTGDFALPVLGVVGTALGLLIIGVNCVIMTVQVKKPLCLQREAKV 300

DB 241 flpmgspspaegstgdfalpvlgvvgtalgliliigvncvimtvqvkklclqreakv 300

QY 301 PHLPADKARGTQGPQOQHLLITAPSSSSLESSASALDRRAPTRNQPAQGVASGAGE 360

DB 301 phlpadkargtgqgpqoqhllitapsssslessasaldraptrnqpqagveasgag 360

QY 361 ARASTGSSDSSPGHGHTQVNVTCIVNVCSSDHSQSSQASSTMGDTDSSPSPKDEQ 420

DB 361 arastgssdsspgghgtqvnvtcivnvcssdhsqssqasstmgdtdsspspkdeq 420

Db 361 arastgssdsspgghgtqvnvtciwvncssdhssqcsqasmtmgdtdsspsespkdq 420  
 QY 421 VPFSEKCAFRLQETPETLLGSTEELKPLPLGVDPAGMKPS 461  
 Db 421 vpfsekecafrsqtletpetllgsteeekplplgvdpagmkps 461

## RESULT 9

AAB37686  
 ID AAB37686 standard; Protein; 461 AA.  
 XX  
 AC AAB37686;  
 XX  
 DT 02-MAR-2001 (first entry)  
 XX  
 DE Human 40 kDa TNF inhibitor precursor.  
 XX  
 XX TNF inhibitor; antiinflammatory; Tumour Necrosis Factor; interleukin;  
 KW IL-1; inflammatory disease; degenerative disease; human; lymphotoxin.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6143866-A.  
 XX  
 PD 07-NOV-2000.  
 XX  
 PF 19-JAN-1995; 95US-0375242.  
 XX  
 PR 19-JUL-1990; 90US-0555274.  
 PR 09-JUL-1993; 93US-0090366.  
 PR 18-JUL-1989; 89US-0381080.  
 PR 11-DEC-1989; 89US-0450329.  
 PR 07-FEB-1990; 90US-0479661.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Squires C, King MW, Hale KK, Brewer MT, Thompson RC;  
 PI Vanderslice RW, Vannice J, Kohno T;  
 XX  
 DR N-PSDB; AAC83951.  
 XX  
 DR Novel 30 kDa tumor necrosis factor inhibitor analog comprising a  
 PT non-native cysteine residue cross-linked with polyethylene glycol,  
 PT useful for treating inflammatory and degenerative diseases mediated by  
 PT TNF -  
 XX  
 XX Example 12; Fig 39; 82pp; English.  
 PS  
 XX The present invention relates to Tumour Necrosis Factor (TNF) inhibitors  
 CC (see AAB37676 and AAB37685), which have TNF inhibitory activity. The  
 CC novel TNF inhibitors of the present invention are useful as therapeutic  
 CC agents for inhibiting the activity of TNF and interleukin (IL-1), and  
 CC for treating inflammatory and degenerative diseases mediated by TNF. The  
 CC present sequence is the precursor for 40 kDa TNF inhibitor. The 40 kDa  
 CC TNF inhibitor can inhibit both TNF alpha and beta (lymphotoxin).  
 XX  
 XX Sequence 461 AA:

Query Match 99.8%; Score 2462; DB 22; Length 461;  
 Best Local Similarity 99.8%; Pred. No. 2.6e-152;  
 Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPVAVWAALVGLLEWAAHALPAQVATPYAPPGSTCRLEYYDQTAQMCCSKSPG 60  
 Db 1 mapvavwaalvglelwaaahalpavqatpyapepgstcrleyydqtacmccskspg 60  
 QY 61 QHAKVCTKTSVTDCSDSTYTLWNWVPECLSCGSRCSDDVETQACTREONRICTC 120  
 Db 61 qhakvctktsdtdvcdscdstytlwnwvpeclscgsrcssddvqetactreqrictc 120  
 QY 121 RPYWCALSKQEGCRLCAPLRKCRPGFGVARGFTETSDVVKPCAPGTFSTNTTSSDIDCR 180

Db 121 rpywcalskqegcrclcaplrkcrpgfgvargftetdsvckpcapgtfnttstdidcr 180  
 QY 181 PHQICNVVAIPGNASMDAVCTSTPTRSMAPGAYHLPPVSTRSOHTQTPTEPSTAPSTS 240  
 Db 181 phqicnvvaipgnasrdavctstptrsmagavhlppqvstrsqhtqtpstapsts 240  
 QY 241 FLPLPMGSPPAEGSTGDFALPVGILIVGTALGLLIIGVWNCVIMTVQVKKKPLCLQREAKV 300  
 Db 241 flplpmgsppeagstgdfaipvgilivgtalglilivwncvimtqvkklplclqreakv 300  
 QY 301 PHLPADKARGTQGPQOHLITAPSSSSSSLESSASALDRRAPTRNOQAPQVEASGAGE 360  
 Db 301 phlpadkargtqgpeqhlitapsssssslessasaldraptrnqpapqveasgag 360  
 QY 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSSSDHSSQCSQASSTMGDTDSSPSKDEQ 420  
 Db 361 arastgssdspgghgtqvnvtciwvncssdhssqcsqasmtmgdtdsspsespkdq 420  
 QY 421 VPFSEKCAFRLQETPETLLGSTEELKPLPLGVDPAGMKPS 461  
 Db 421 vpfsekecafrsqtletpetllgsteeekplplgvdpagmkps 461

## RESULT 10

AAR72504  
 ID AAR72504 standard; Protein; 461 AA.  
 XX  
 AC AAR72504;  
 XX  
 DT 31-OCT-1995 (first entry)  
 XX  
 DE p75 Tumour Necrosis Factor Receptor.  
 XX  
 KW Ligand; tumour necrosis factor; nerve growth factor; TNF; NGF;  
 KW receptor.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 27..214  
 FT /label= TBP11.  
 FT Domain 258..285  
 FT /label= Transmembrane domain.  
 FT Misc-difference 259  
 FT /note= "Unidentified amino acid."  
 XX  
 PN EP648783-A.  
 XX  
 PD 19-APR-1995.  
 XX  
 PF 11-OCT-1994; 94EP-0116015.  
 XX  
 PR 12-OCT-1993; 93IL-0107267.  
 XX  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 PA (WALL/) WALLACH D.  
 XX  
 PI Beletsky I, Bigda J, Mett I, Wallach D;  
 XX  
 DR WPI; 1995-148673/20.  
 DR N-PSDB; AAQ89544.  
 XX  
 PT Tumour necrosis factor (TNF) receptor ligand - used to increase  
 PT inhibitory effect of a soluble TNF receptor  
 PS  
 PS Disclosure; Figure 2; 18pp; English.  
 XX  
 CC A ligand to a member of the tumour necrosis factor (TNF)/nerve  
 CC growth factor (NGF) receptor family which binds either to the region  
 CC of the 4th-Cys rich domain of the receptor, or to the region between  
 CC it and the cell membrane may be used in the production of a  
 CC pharmaceutical composition for increasing the inhibitory effect of a

```

CC soluble receptor of the TNF/NGF receptor family. This sequence
CC is the sequence of the p75 TNF receptor.
XX
SQ Sequence 461 AA;

Query Match 97.0%; Score 2394; DB 16; Length 461;
Best Local Similarity 96.4%; Pred. No. 6.7e-148;
Matches 451; Conservative 0; Mismatches 3; Indels 14; Gaps 2;

QY 1 MAPVAVWAALAVGLELWAAHALPAQVAFTPYAPEPGSTCRLREYYDQTAQMCCKSPG 60
Db 1 mapvavwaalavglelwaahalpaqvaftpyapepgstcrlreydqtaqmcckspg 60
QY 61 QHAKVFCTKTSDTVCDCEDSTYTLQNNWVPECLSCGRSSDQVETQACTREQNRICTC 120
Db 61 qhakvfctktsdvtcdcedstytqlnnwvpeclscgrssdqvetaactreqnrictc 120
QY 121 RPYGWCALSKOEGCRLCAPLRKCRPGFVGARPGTETSDVVCCKPCAPGTFSTSTDICR 180
Db 121 rpygwcalskqegcrlcaplrkcrpgfvgarpgtetsdvvcckpcapgtfststsdicr 180
QY 181 PHQICNVVAIPGNASMDAVCTSTPTRSMAPCAVHLPQVSTRSQHTQPTPEPSTAPSTS 240
Db 181 phqicnvvaipgnasmdavctstptrsmapcavhlpqvpvstrsqhtqptpepstapsts 240
QY 241 FLLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNCVIMTVQKKPLCLQREAKV 300
Db 241 flpmpgspgaegstgdfalpvglivgtalgliliigvncvmtvqkkplclqreakv 300
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPAQVGEASGAGE 360
Db 301 phlpadkargtqgpeqhllitapsssssslessasaldrraptrnqpapqveasgag 360
QY 361 ARASTGSSDSPGGHGTHQ-----VNVTCIVNVCSDDHSSQCSSQASSTMGDTDSSES 413
Db 361 arastgssdspgghgtqapgvaeasvntctvnnvcsddhssqcssqassumgtdssps 413
QY 414 ESPKDEOVPFSEKCAPRSQLETPETLGSTEERKPLPLGVDPDAGMKPS 461
Db 414 espkdeovpfsekecaprsqletpetllgsteekpplpvgpdmgmkps 461

RESULT 11
AAR51002
ID AAR51002 standard; Protein; 461 AA.
XX
AC AAR51002;
XX
XX 07-OCT-1994 (first entry)
XX
DE Sequence of human tumour necrosis factor receptor type I
DE (TNFRI).
XX
XX Tumour necrosis factor receptor; type I; TNFRI; arthritis therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal
FT Peptide 22..461
FT /label= mature
XX
XX W09406476-A.
XX
XX 31-MAR-1994.
XX
XX 14-SEP-1993; 93WO-US08666.
XX
XX 15-SEP-1992; 92US-0946236.
XX
XX (IMMUNEX ) IMMUNEX CORP.
XX
XX

```

---

```

XX Jacobs CA, Smith CA;
XX WPI; 1994-118172/14.
XX N-PSDB; AQA45224.
XX
XX Treating TNF mediated inflammatory diseases with TNF antagonist -
XX esp. soluble form of TNF receptor, opt. as fusion protein with
XX human immunoglobulin Fc region, esp. for treating arthritis
XX
XX Disclosure; Page 28-30; 47pp; English.
XX
XX AQA45224 is cDNA from clone 1 of library WI-26 VA4 of human fibroblast
XX cell line WI-26 VA4. The mature full-length TNFRI is a glycoprotein
XX having a mol. wt. of about 75-80 kDa. The cloning of the cDNA for
XX TNFRI was described in Smith et al., Science 248:1019, 1990. Clone 1
XX is contained in expression vector pCAV/NOT-TNFR (ATCC 68088). The
XX preferred TNFRs of the present invention are soluble forms of TNFRI
XX and TNFRII having at least 20 AAs. Soluble TNFR constructs are
XX devoid of a transmembrane region but retain the ability to bind TNF.
XX Examples of soluble TNFRs are huTNFRdelta235, huTNFRdelta185 and
XX huTNFRdelta163 which encode respectively AAs 1-235, 1-185 and 1-163
XX of AAR51002. An equivalent soluble TNFR is huTNFRdelta x wherein x
XX is selected from any one of AAs 163-235 of AAR51002.
XX
XX Sequence 461 AA;

Query Match 96.3%; Score 2376; DB 15; Length 461;
Best Local Similarity 96.7%; Pred. No. 9.9e-147;
Matches 446; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAPVAVWAALAVGLELWAAHALPAQVAFTPYAPEPGSTCRLREYYDQTAQMCCKSPG 60
Db 1 mapvavwaalavglelwaahalpaqvaftpyapepgstcrlreydyqtaemcwkscspg 60
QY 61 QHAKVFCTKTSDTVCDCEDSTYTLQNNWVPECLSCGRSSDQVETQACTREQNRICTC 120
Db 61 qhakvfctktsdvtcdcedstytqlnnwvpeclscgrssdqvetaactreqnrictc 120
QY 121 RPYGWCALSKOEGCRLCAPLRKCRPGFVGARPGTETSDVVCCKPCAPGTFSTSTDICR 180
Db 121 rpygwcalskqegcrlcaplrkcrpgfvgarpgtetsdvvcckpcapgtfststsdicr 180
QY 181 PHQICNVVAIPGNASMDAVCTSTPTRSMAPCAVHLPQVSTRSQHTQPTPEPSTAPSTS 240
Db 181 phqicnvvaipgnasmdavctstptrsmapcavhlpqvpvstrsqhtqptpepstapsts 240
QY 241 FLLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNCVIMTVQKKPLCLQREAKV 300
Db 241 flpmpgspgaegstgdfalpvglivgtalgliliigvncvmtvqkkplclqreakv 300
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPAQVGEASGAGE 360
Db 301 phlpadkargtqgpeqhllitapsssssslessasaldrraptrnqpapqveasgag 360
QY 361 ARASTGSSDSPGGHGTHQVNVTCIVNVCSDDHSSQCSSQASSTMGDTDSSESFPKDDQ 420
Db 361 arastgssdspgghgtqvnvtcivnvcsddhssqcssqassmgtdtdsspsespkdq 420
QY 421 VPFSKECAFBSQLETPETLGSTEERKPLPLGVDPDAGMKPS 461
Db 421 vpfskecafbsqletpetllgsteekpplpvgpdmgmkps 461

RESULT 12
AAY30935
ID AAY30935 standard; Protein; 392 AA.
XX
XX AAY30935;
XX
XX 18-OCT-1999 (first entry)
XX

```

DE Human tumour necrosis factor binding protein fragment.

XX Tumour necrosis factor binding protein; TNF; insoluble protein; agonist;  
KW anti-inflammatory; antimalarial; treatment; septic shock; inflammation;  
KW autoimmune glomerulonephritis; cerebral malaria; immune response;  
KW antagonist; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FH Protein 1..392

FT /note= "Partial sequence, no start codon given"

XX EP939121-A2.

XX 01-SEP-1999.

XX 31-AUG-1990; 90EP-0116707.

XX 20-APR-1990; 90CH-0001347.

PR 12-SEP-1989; 89CH-0003319.

PR 08-MAR-1990; 90CH-0000746.

XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX Brockhaus M, Dembic Z, Gentz R, Lesslauer W, Loetscher H;

PI Schlaeger E;

XX WPI; 1999-480840/41.

DR N-PSDB; AAZ09171.

XX New insoluble proteins, and fragments, that bind to tumor necrosis

PT factor, used to treat e.g. septic shock or cerebral malaria

XX Claim 4a; Fig 4; 25pp; German.

PS This invention describes novel homogeneous insoluble proteins (I),  
CC their (in)soluble fragments (Ia) and their salts that can bind tumour  
CC necrosis factor (TNF). The products of the invention have  
CC anti-inflammatory and antimalarial activity. (i) and (Ia) are used (i)  
CC to treat diseases in which TNF is involved (e.g. septic shock, autoimmune  
CC glomerulonephritis, cerebral malaria, immune responses and inflammation),  
CC (ii) to purify TNF, (iii) to identify TNF (antagonists and (iv) for  
CC diagnostic determination of TNF in body fluids. Antibodies raised against  
CC (I) are used for affinity purification of (I). This sequence represents  
CC a tumour necrosis factor binding protein fragment described in the method  
CC of the invention.

XX Sequence 392 AA;

Query Match 83.1%; Score 2051.5; DB 20; Length 392;

Best Local Similarity 98.7%; Pred. No. 9.8e-126;

Matches 387; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 71 SDTVCDSCDSTYTQIWNVPECLSCGSCSSDQVETQACTREONRICRPGWYCALSK 130

DB 1 sdsdvcscdsstytqlwnvpeclscgscssdqvetqactreqnricrpgwycalsk 60

QY 131 QEGCRLCAPLRCRPGFGVARGPTETSDVVKPCAPGTFSTTSDICRPHQICNVVAI 190

DB 61 qegcrlcaplrcrpgfgvargptetdvckpcapgtfsttsdtdicrphqicnvai 120

QY 191 PCNASMDAVCTSTPSTRMAPCAVHLPPQVSTRSQTPTPEPSTAPSTFLLPMGSPSP 250

DB 121 pgnasrdavctstprsmagavhlppqvstrsqtptpspstapstsfllpmgspsp 180

QY 251 AEGSGDFALPGLVIGVTALGLLIIGVNVCVIMTQVKKKPLCLQREAKVPHLPADKARG 310

DB 181 aegsgdfalpgvlvigtalglilggnvncvmtqvkklpcklqreakvphlpadkarg 240

QY 311 TQGPQQHLLITAPSSSSSSLESSALSALDRRAPTRNQOPAGVEASGAGEARASTGSS-D 369

DB 1 tgppeqqhllitapsssssslessalsaldraptrnqpqpgvveasgagearastgssad 300

Db 241 tqgpeqqhllitapsssssslessalsaldraptrnqpqpgvveasgagearastgssad 300

QY 370 SSGGCGTQVNVTCIVNVCSHSSQCSQASSTMGDTSSPSESPPKDEOVPSKEBCA 429

Db 301 sspgghgtqvnvtcivnvcsdhssqcsqasstmgdtsspsespdkedqvfskeca 360

QY 430 FRSQLETPETLLGSTERKPLPLGVDPDAGMKPS 461

Db 361 frsqletpetllgsteeekplplgvdpdagmkps 392

RESULT 13

AAR11605

ID AAR11605 standard; Protein; 392 AA.

XX AAR11605;

XX 24-MAY-1991 (first entry)

XX Human 75kD TNF-binding protein.

XX Tumour Necrosis Factor; binding proteins; septic shock;

XX autoimmune glomerulonephritis; lymphokine; cytokine.

XX Key Location/Qualifiers

FH Misc-difference 3 /label= Ser, Thr

FT EP417563-A.

XX 20-MAR-1991.

XX 31-AUG-1990; 90EP-0116707.

XX 20-APR-1990; 90CH-0001347.

PR 12-SEP-1989; 89CH-0003319.

PR 08-MAR-1990; 90CH-0000746.

XX (HOFF ) HOFFMANN-LA ROCHE AG.

XX Brockhaus M, Dembic Z, Gentz R, Lesslauer W, Loetscher H;

PI Schlaeger EJ;

XX WPI; 1991-081851/12.

DR N-PSDB; AAQ10956.

XX Insoluble tumour necrosis factor binding proteins - and DNA

PT encoding them, useful in pharmaceutical prods. and for antibody

PT prodn.

XX Claim 1; Fig 1; 26pp; German.

XX Partial amino acid sequences were determined for the 55 and 75kD

CC TNF-BPs (see AAR1072-R11081) and oligonucleotide primers were used

CC synthesised based on these partial sequences. The primers were used

CC to produce a cDNA fragment for use as a probe to screen a human

CC placental cDNA bank constructed in lambda gt11. Positive clones were

CC identified and sequenced. Repeated sequencing showed a discrepancy

CC at position 7 such that the third codon encodes either Thr or Ser.

CC See also AAQ10955.

XX Sequence 392 AA;

Query Match 83.1%; Score 2050.5; DB 12; Length 392;

Best Local Similarity 98.7%; Pred. No. 1.1e-125;

Matches 387; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 71 SDTVCDSCDSTYTQIWNVPECLSCGSCSSDQVETQACTREONRICRPGWYCALSK 130

DB 1 sdsdvcscdsstytqlwnvpeclscgscssdqvetqactreqnricrpgwycalsk 60

QY 131 QEGCRLCAPLRCRPGFGVARGPTETSDVVKPCAPGTFSTTSDICRPHQICNVVAI 190



```

Db 61 qegrlcaplpcrpgfgrvrgtetsdvckpcapgtfisttsdtdcrphqicnvai 120
QY 191 PGNASMDAVCTSTPTRSMAPGAVHLPPQVSTRSOHTOPTPEPSTAPSTSLPLPMGSPSP 250
Db 121 pgnasrdavctstptrsmagavhlppqvstrsqhtqpspepstapstslplpmgspsp 180
QY 251 AEGSTGDFALPVLGVLGVTALGLLIIGVNCVIMTOVKKPLCLQREAKVPHLPADKARG 310
Db 181 aegstgdfalpvglvgvtalglilgvcvntvmtqvkklpqlclqreakvphlpadkarg 240
QY 311 TQPEQOHLITAPSSSSSSSSSSSASALDRAPTRNQPQPGVPSGAGEARASTGSS-D 369
Db 241 tqpeqghlilitapsssssssslessasaldrraptrnqpqpgvpsgagearastgssad 300
QY 370 SSPGGHCTQVNTCIVNVCSSSHSSSCQASSTMGDTSSPSKDEQVPFSKECA 429
Db 301 sspgghtqvnvtcivnvcssdhssscqasstmgdtdsspskdeqvptskeca 360
QY 430 FRSQLETPETLLGSTEELKPLPLGVDPDAGMKPS 461
Db 361 frsqletpetllgsteekplplgvdpdagmkps 392

```

## RESULT 14

AAR11142  
ID AAR11142 standard; Protein; 474 AA.

AC AAR11142;

XX 24-MAY-1991 (first entry)

DE TNF-R deduced from TNF-R clone 11.

XX Tumour necrosis factor receptor; immune response; inflammation;

KW cachexia; septic shock.

XX Mus musculus strain C57BL/6.

OS Key Location/Qualifiers

FH Peptide 1..22

FT /label= signal sequence

FT 234..265

FT /label= transmembrane region

XX EP418014-A.

XX 20-MAR-1991.

PD 10-SEP-1990; 90EP-0309875.

XX 10-MAY-1990; 90US-0523635.

PR 11-SEP-1989; 89US-0405370.

PR 13-OCT-1989; 89US-0421417.

XX (IMMU-) IMMUNEX CORP.

XX Smith CA, Goodwin RG, Beckmann PM;

XX WPI; 1991-082230/12.

DR N-PSDB; AAQ10991.

XX New tumour necrosis factor -alpha and -beta receptors - and DNA

PT encoding these used to regulate immune responses in treatment of

PT cachexia, septic shock or side-effects of cytokine therapy.

XX Disclosure; Fig 3; 41pp; English.

XX The sequence was deduced from a clone isolated from library prepd.

CC from a murine T helper cell line, 7B9.

CC See also AAR11141.

XX Sequence 474 AA;

SQ

Query Match 61.0%; Score 1506; DB 12; Length 474;

Best Local Similarity 62.8%; Pred. No. 3.3e-90;

Matches 292; Conservative 50; Mismatches 115; Indels 8; Gaps 6;

```

QY 1 MAPVAVMAALAVGLLEMAAALPAQVAFTPYAPGPGSTCRL-REYYDOTAQMCCSKCSP 59
Db 1 mapaalwvalvelfqlwatghtvpaqvvltpykpepgyecqisqeyydrkqmcckakcp 60
QY 60 QOHAKVCTKTSDIVCDSCEDSTYTQLNWNVPECLSCSCSSDQVETQACTREQNRICT 119
Db 61 ggyvkhfckntsdtdvcaadceasmtyqwnqfrtclscsscttdgveiractkqgnrvca 120
QY 120 CRPGWYCALSKQEG-CRLCAPLRKCRPGVGVARPCTETSDVVCKPCAPGTENTTSSTD 178
Db 121 ceagrycdlkhsgscrcqmriskgpgfvgvassrapngvnickacaptgfdtsdstdv 180
QY 179 CRPHQICNVVAIPGNASMDVCTSTPTRSMAPGAVHLPPQVSTRSOHTOPTPEPSTAPS 238
Db 181 crphricsailaipgnastdvcapesptlsalprtlvysqpeptrsqpldgpgpsqtp- 239
QY 239 TSFLLPMGSPPAEGST-GDFALPVLGVLGVTALGLLIIGVNCVIMTOVKKPLCLQRE 297
Db 240 -siltsgstpliegskgglplglilvgvtsigllmglivncilvqrkkksclgrd 298
QY 298 AKVPHLPADKARGTQGPQOHLITAPSSSSSSSSSALDRRAPTRNQPQPGV-EAS 356
Db 299 akvphvpdeksgdavgleqqhlittapsssssslessasaldrraptrnqpqpgvmae 358
QY 357 GAGEARASTGSSDSSPGGHGTQVNVTCIVNVCSSSHSSSCQASSTMGDTSSPSSESP 416
Db 359 gfgearassrisdsdshgshgthvntcivnvcssdhssscqasatvgdpdakpsasp 418
QY 417 KDEQVPFSKECAPRSQLETPETLLGSTEELKPLPLGVDPDAGMKPS 461
Db 419 kdeqvfpfsgeecpsqpcettetl--qshekplplgvdpdmgmks 461

```

## RESULT 15

AAB70001  
ID AAB70001 standard; protein; 518 AA.

AC AAB70001;

XX 08-MAY-2001 (first entry)

DT STNFR(075):Fc fusion protein.

DE TNFR; tumour necrosis factor receptor; immunoglobulin constant region;

XX FC; STNFR(075):Fc; fusion protein; gene therapy; arthritis;

XX recombinant adeno-associated virus; rAAV; antiarthritic.

XX Unidentified.

XX WO200111034-A2.

XX 15-FEB-2001.

XX 08-AUG-2000; 2000WO-US21693.

XX 09-AUG-1999; 99US-0160080.

XX (TARG-) TARGETED GENETICS CORP.

XX Carter B;

XX WPI; 2001-191537/19.

XX New recombinant viral vectors comprising a single-stranded heterologous

PT nucleotide sequence have a region that forms intrastrand base pairing,

PT useful in gene therapy and genomics screening

XX

Search completed: February 12, 2002, 12:47:38  
Job time: 160 sec

GenCore version 4.5  
Copyright (c) 1993 -2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:45:28 ; Search time 15.87 Seconds  
(without alignments)  
653.687 Million cell updates/sec

Title: US-09-800-909-2  
Perfect score: 2468  
Sequence: 1 MAPVAVWAALAVGLELWAAA.....GSTEEKPLPLGVDPAGMKPS 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2468	100.0	461	1 US-08-385-229-2	Sequence 2, Appli
2	2468	100.0	461	2 US-08-650-000-2	Sequence 2, Appli
3	2468	100.0	461	4 US-08-477-347-3	Sequence 3, Appli
4	2468	100.0	461	4 US-08-476-862-2	Sequence 2, Appli
5	2468	100.0	461	6 5395760-2	Patent No. 5395760
6	2462	99.8	461	4 US-09-042-785A-7	Sequence 7, Appli
7	2462	99.8	461	4 US-09-006-353A-4	Sequence 4, Appli
8	1512	61.3	474	2 US-08-650-000-4	Sequence 4, Appli
9	1512	61.3	474	4 US-09-042-785A-8	Sequence 8, Appli
10	1512	61.3	474	6 5395760-4	Patent No. 5395760
11	1433	58.1	518	1 US-08-385-229-4	Sequence 4, Appli
12	1404	56.9	486	1 US-08-243-010-1	Sequence 1, Appli
13	1312	53.2	235	3 US-09-326-394-4	Sequence 4, Appli
14	1263	51.2	227	3 US-08-974-022-48	Sequence 48, Appli
15	1263	51.2	227	4 US-08-795-445A-48	Sequence 48, Appli
16	1263	51.2	227	4 US-08-795-447A-48	Sequence 48, Appli
17	1263	51.2	227	4 US-08-974-186-48	Sequence 48, Appli
18	1263	51.2	227	4 US-08-795-446B-48	Sequence 48, Appli
19	931	37.7	163	2 US-08-219-237B-5	Sequence 5, Appli
20	931	37.7	163	4 US-08-477-347-13	Sequence 13, Appli
21	931	37.7	163	4 US-08-476-862-4	Sequence 4, Appli
22	931	37.7	163	4 US-08-468-560C-5	Sequence 5, Appli
23	924.5	37.5	164	2 US-08-232-087A-9	Sequence 9, Appli
24	695	28.2	120	3 US-08-974-022-42	Sequence 42, Appli
25	695	28.2	120	4 US-08-795-445A-42	Sequence 42, Appli
26	695	28.2	120	4 US-08-795-447A-42	Sequence 42, Appli
27	695	28.2	120	4 US-08-974-186-42	Sequence 42, Appli

28 695 28.2 120 4 US-08-795-446B-42 Sequence 42, Appli  
29 453 18.4 77 4 US-08-866-545-2 Sequence 2, Appli  
30 379.5 15.4 349 4 US-09-006-353A-13 Sequence 13, Appli  
31 373 15.1 355 1 US-08-292-549-6 Sequence 6, Appli  
32 373 15.1 355 4 US-09-006-353A-14 Sequence 14, Appli  
33 360.5 14.6 326 1 US-08-292-549-4 Sequence 4, Appli  
34 360.5 14.6 326 5 PCT-US91-02207-4 Sequence 4, Appli  
35 351.5 14.2 300 2 US-08-794-796-2 Sequence 2, Appli  
36 343.5 13.9 605 4 US-09-042-785A-23 Sequence 23, Appli  
37 343.5 13.9 655 3 US-08-959-382-2 Sequence 2, Appli  
38 342 13.9 299 4 US-09-286-529-17 Sequence 17, Appli  
39 340 13.8 211 4 US-09-286-529-20 Sequence 20, Appli  
40 330 13.4 401 3 US-08-974-022-6 Sequence 6, Appli  
41 330 13.4 401 4 US-09-042-785A-12 Sequence 12, Appli  
42 330 13.4 401 4 US-08-795-445A-6 Sequence 6, Appli  
43 330 13.4 401 4 US-08-795-447A-6 Sequence 6, Appli  
44 330 13.4 401 4 US-08-974-186-6 Sequence 6, Appli  
45 330 13.4 401 4 US-08-795-446B-6 Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-08-385-229-2  
; Sequence 2, Application US/08385229  
; Patent No. 5605690  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Cindy A.  
; APPLICANT: Smith, Craig A.  
; TITLE OF INVENTION: Method of Treating TNF-Dependent  
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/385,229  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/946,236  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wright, Christopher L.  
; REGISTRATION NUMBER: 31,680  
; REFERENCE/DOCKET NUMBER: 2503  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 587-0606  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 461 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-385-229-2

Query Match 100.0%; Score 2468; DB 1; Length 461;  
Best Local Similarity 100.0%; Pred. No. 4.2e-173;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPVAVWAALAVGLELWAAAHALPAQVAFTPYAPEPGSTCRRLREYDQTAQMCCSKSPG 60

Db 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPEPGSTCRLREYDQTAQMCCSKCSPG 60  
QY 61 QHAKVFCTKTSDDTCDSCEDSTYTLQNNWVPECLSCGSRSSDOVETQACTREONR1CTC 120  
Db 61 QHAKVFCTKTSDDTCDSCEDSTYTLQNNWVPECLSCGSRSSDOVETQACTREONR1CTC 120  
QY 121 RPYGMYCALSKEGCRCLCAPLRCRPGFVARPGTETSDVVKCKPCAPGTFSTNTSTDICR 180  
Db 121 RPYGMYCALSKEGCRCLCAPLRCRPGFVARPGTETSDVVKCKPCAPGTFSTNTSTDICR 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
Db 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
QY 241 FLPLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNVCMVIMTVQVKKPLCLQREAKV 300  
Db 241 FLPLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNVCMVIMTVQVKKPLCLQREAKV 300  
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPQAPGVEASGAGE 360  
Db 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPQAPGVEASGAGE 360  
QY 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSDDHSSQCSQASSTMGDTDSSPSPKDEQ 420  
Db 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSDDHSSQCSQASSTMGDTDSSPSPKDEQ 420  
QY 421 VPFKSKECAFRSOLETPETLLGSTEERKPLPLGVDPAGMKPS 461  
Db 421 VPFKSKECAFRSOLETPETLLGSTEERKPLPLGVDPAGMKPS 461

RESULT 2  
US-08-650-000-2  
; Sequence 2, Application US/08650000  
; Patent No. 5945397  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Craig A.  
; APPLICANT: Goodwin, Raymond G.  
; APPLICANT: Beckmann, M. Patricia  
; TITLE OF INVENTION: Tumor Necrosis Factor Receptors  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/650,000  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,453  
; FILING DATE:  
; APPLICATION NUMBER: US/08/039,765  
; FILING DATE:  
; APPLICATION NUMBER: US 403,241  
; FILING DATE: 05-SEP-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 405,370  
; FILING DATE: 11-SEP-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 421,417  
; FILING DATE: 13-OCT-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 523,635

; FILING DATE: 10-MAY-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wight, Christopher L.  
; REGISTRATION NUMBER: 31,680  
; REFERENCE/DOCKET NUMBER: 2501-D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 461 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-650-000-2  
  
Query Match 100.0%; Score 2468; DB 2; Length 461;  
Best Local Similarity 100.0%; Pred. No. 4.2e-173;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPEPGSTCRLREYDQTAQMCCSKCSPG 60  
Db 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPEPGSTCRLREYDQTAQMCCSKCSPG 60  
QY 61 QHAKVFCTKTSDDTCDSCEDSTYTLQNNWVPECLSCGSRSSDOVETQACTREONR1CTC 120  
Db 61 QHAKVFCTKTSDDTCDSCEDSTYTLQNNWVPECLSCGSRSSDOVETQACTREONR1CTC 120  
QY 121 RPYGMYCALSKEGCRCLCAPLRCRPGFVARPGTETSDVVKCKPCAPGTFSTNTSTDICR 180  
Db 121 RPYGMYCALSKEGCRCLCAPLRCRPGFVARPGTETSDVVKCKPCAPGTFSTNTSTDICR 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
Db 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHTQPTPEPSTAPSTS 240  
QY 241 FLPLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNVCMVIMTVQVKKPLCLQREAKV 300  
Db 241 FLPLPMGSPPAEGSTGDFALPVGLIVGTALGLLIIGVNVCMVIMTVQVKKPLCLQREAKV 300  
QY 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPQAPGVEASGAGE 360  
Db 301 PHLPADKARGTQGPQEQHLLITAPSSSSSSLESSASALDRRAPTRNQPQAPGVEASGAGE 360  
QY 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSDDHSSQCSQASSTMGDTDSSPSPKDEQ 420  
Db 361 ARASTGSSDSPGGHGTQVNVTCIVNVCSDDHSSQCSQASSTMGDTDSSPSPKDEQ 420  
QY 421 VPFKSKECAFRSOLETPETLLGSTEERKPLPLGVDPAGMKPS 461  
Db 421 VPFKSKECAFRSOLETPETLLGSTEERKPLPLGVDPAGMKPS 461

RESULT 3  
US-08-477-347-3  
; Sequence 3, Application US/08477347  
; Patent No. 6232446  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: BIGDA, Jacek  
; APPLICANT: BELETSKY, Igor  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: TNF LIGANDS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:



```
Db 241 FLPMGSPSPAEGSTGDFALPVGLIIVGTALGILLIIGVNCVINTQVKKKPLCLQREAKV 300
QY 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
Db 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
QY 361 ARASTSSDSSPGGHGTQVNVTCIVNVCSSDSSDHSSQCSSQASSTMGDTDSSPSPKDEQ 420
Db 361 ARASTSSDSSPGGHGTQVNVTCIVNVCSSDSSDHSSQCSSQASSTMGDTDSSPSPKDEQ 420
QY 421 VPFSKEECAFRSOLETPETLLGSTEKPLPLGVPDAGMKPS 461
Db 421 VPFSKEECAFRSOLETPETLLGSTEKPLPLGVPDAGMKPS 461

RESULT 5
5395760-2
: Patent No. 5395760
: APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN,
: M. PATRICIA
: TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
: B-RECEPTORS
: NUMBER OF SEQUENCES: 17
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/523,635
: FILING DATE: 10-MAY-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 421,417
: FILING DATE: 13-OCT-1989
: APPLICATION NUMBER: 405,370
: FILING DATE: 11-SEP-1989
: APPLICATION NUMBER: 403,241
: FILING DATE: 05-SEP-1989
: SEQ ID NO:2:
: LENGTH: 461
5395760-2

Query Match 100.0%; Score 2468; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 4.2e-173;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLREYDQTQAMCCSKSPG 60
Db 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLREYDQTQAMCCSKSPG 60
QY 61 QHAKVFCTKTSDTVDCSDCEDSTYTQLWNWVPECLSCGSRSSDQVETQACTREQNRICTC 120
Db 61 QHAKVFCTKTSDTVDCSDCEDSTYTQLWNWVPECLSCGSRSSDQVETQACTREQNRICTC 120
QY 121 RPYWCALSKQEGCRLCAPLRCRPGFVGVARPGTETSDVVKPCAPGTFSTNTSSSTDICR 180
Db 121 RPYWCALSKQEGCRLCAPLRCRPGFVGVARPGTETSDVVKPCAPGTFSTNTSSSTDICR 180
QY 181 PHOICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPPQPVSTRSQHTQPTPEPSTAPSTS 240
Db 181 PHOICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPPQPVSTRSQHTQPTPEPSTAPSTS 240
QY 241 FLPMGSPSPAEGSTGDFALPVGLIIVGTALGILLIIGVNCVINTQVKKKPLCLQREAKV 300
Db 241 FLPMGSPSPAEGSTGDFALPVGLIIVGTALGILLIIGVNCVINTQVKKKPLCLQREAKV 300
QY 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
Db 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
QY 361 ARASTSSDSSPGGHGTQVNVTCIVNVCSSDSSDHSSQCSSQASSTMGDTDSSPSPKDEQ 420
Db 361 ARASTSSDSSPGGHGTQVNVTCIVNVCSSDSSDHSSQCSSQASSTMGDTDSSPSPKDEQ 420
QY 421 VPFSKEECAFRSOLETPETLLGSTEKPLPLGVPDAGMKPS 461
Db 421 VPFSKEECAFRSOLETPETLLGSTEKPLPLGVPDAGMKPS 461
```

```
RESULT 6
US-09-042-785A-7
: Sequence 7, Application US/09042785A
: Patent No. 6194151
: GENERAL INFORMATION:
: APPLICANT: Busfield, Samantha J
: TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
: AND USES THEREFOR
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/042.785A
: FILING DATE: 17-MAR-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/938,896
: FILING DATE: 26-SEP-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragouras, Amy E
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: MEI-001CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400.
: TELEFAX: (617)742-4214
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 461 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: internal
US-09-042-785A-7
```

```
Query Match 99.8%; Score 2462; DB 4; Length 461;
Best Local Similarity 99.8%; Pred. No. 1.2e-172;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLREYDQTQAMCCSKSPG 60
Db 1 MAPVAVWAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTCRLREYDQTQAMCCSKSPG 60
QY 61 QHAKVFCTKTSDTVDCSDCEDSTYTQLWNWVPECLSCGSRSSDQVETQACTREQNRICTC 120
Db 61 QHAKVFCTKTSDTVDCSDCEDSTYTQLWNWVPECLSCGSRSSDQVETQACTREQNRICTC 120
QY 121 RPYWCALSKQEGCRLCAPLRCRPGFVGVARPGTETSDVVKPCAPGTFSTNTSSSTDICR 180
Db 121 RPYWCALSKQEGCRLCAPLRCRPGFVGVARPGTETSDVVKPCAPGTFSTNTSSSTDICR 180
QY 181 PHOICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPPQPVSTRSQHTQPTPEPSTAPSTS 240
Db 181 PHOICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPPQPVSTRSQHTQPTPEPSTAPSTS 240
QY 241 FLPMGSPSPAEGSTGDFALPVGLIIVGTALGILLIIGVNCVINTQVKKKPLCLQREAKV 300
Db 241 FLPMGSPSPAEGSTGDFALPVGLIIVGTALGILLIIGVNCVINTQVKKKPLCLQREAKV 300
QY 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
Db 301 PHLPADKARGTQGEQOHLITAPSSSSSSLESSASALDRRAPTRNQPAPGVEASGAGE 360
```

QY	361	ARASTCSSDSSPGGHGTQVNTVCIVNVCSDDHSSQSSQASSTMGDTDSSPSSEPKDEQ	420
Db	361	ARASTCSSDSSPGGHGTQVNTVCIVNVCSDDHSSQSSQASSTMGDTDSSPSSEPKDEQ	420
QY	421	VPSKECAFRSOLETPETLLGSTGTEKPLPLGVPDAGMKPS	461
Db	421	VPSKECAFRSOLETPETLLGSTGTEKPLPLGVPDAGMKPS	461
<p>RESULT 7</p> <p>US-09-006-353A-4</p> <p>; Sequence 4, Application US/09006353A</p> <p>; Patent No. 6261801</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: WEL, YING-FEI</p> <p>; APPLICANT: YU, GUO-LIANG</p> <p>; APPLICANT: GENTZ, REINER</p> <p>; APPLICANT: RUBEN, STEVEN</p> <p>; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5</p> <p>; NUMBER OF SEQUENCES: 26</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: HUMAN GENOME SCIENCES, INC.</p> <p>; STREET: 9410 KEY WEST AVENUE</p> <p>; CITY: ROCKVILLE</p> <p>; STATE: MD</p> <p>; COUNTRY: US</p> <p>; ZIP: 20850</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC Compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/09/006,353A</p> <p>; FILING DATE:</p> <p>; CLASSIFICATION: 435</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: BROOKES, ANDERS A.</p> <p>; REGISTRATION NUMBER: 36,373</p> <p>; REFERENCE/DOCKET NUMBER: PF341</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (301) 309-8504</p> <p>; TELEFAX: (301) 309-8512</p> <p>; INFORMATION FOR SEQ ID NO: 4:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 461 amino acids</p> <p>; TYPE: amino acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: protein</p> <p>US-09-006-353A-4</p>			
<p>Query Match 99.8%; Score 2462; DB 4; Length 461;</p> <p>Best Local Similarity 99.8%; Pred. No. 1.2e-172;</p> <p>Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps</p>			
QY	1	MAPVAVMAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTGRLREYYDQTAQMCCSKCSPG	60
Db	1	MAPVAVMAALAVGLELWAAHAHALPAQVAFTPYAPEPGSTGRLREYYDQTAQMCCSKCSPG	60
QY	61	QHAKVCTKTSDFVCDSCEDSTYTQLWNWVPECLSCGRCSSDQVETQACTRQNRICTC	120
Db	61	QHAKVCTKTSDFVCDSCEDSTYTQLWNWVPECLSCGRCSSDQVETQACTRQNRICTC	120
QY	121	RPGWYCALSKQEGCRLCAPLRKCRPGFGVARGPCTETSDVVCVKPCAPGTFSTNTSSDIDICR	180
Db	121	RPGWYCALSKQEGCRLCAPLRKCRPGFGVARGPCTETSDVVCVKPCAPGTFSTNTSSDIDICR	180
QY	181	PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPPQVSTRSQHTQPTPEPSTAPSTS	240
Db	181	PHQICNVVAIPGNASMDVACTSTSPTRSMAPGAVHLPPQVSTRSQHTQPTPEPSTAPSTS	240





Matches 293; Conservative 50; Mismatches 114; Indels 8; Gaps 6;

QY 1 MAPVAVMAALAVGLELWAAHALPAQVAFYAPPEGSCRL-REYDDQTAQCCSKCSP 59  
DB 1 MAPAAUWALVFLQJLWATHTVPAQVLTTPYKPEGYEQISOEYDRAQCAKCPP 60  
QY 60 GQAKVCTKTSYDVCDCEDSTYTQLMNWPCLSCGSCSDQVETQACTREQNRIC 119  
DB 61 GQYVHFCKNTSDTVCADEASMYTQVWQFRTCLSCSSCTTDOVEIRACTQOQNRVCA 120  
QY 120 CRPGWCYALSKQBG-CRLCAPLRKCRPGFVARPGTETSDVCKPCAPGTFSTTSSTDI 178  
DB 121 CEAGRYCALKTHSGCRQCRLSKCGPGFVASSRAPNGNVLCACAPGFSDTTSSTDV 180  
QY 179 CRPHQCNVVAIPGNASMDAVCTSTPTSRMAPCAVHLPOPVSTRSQHTOPTPEPSTAPS 238  
DB 181 CRPHRCISLAIIPGNASTDAVAPETSLAIPRTLYVSQPEPTRSQPLDOECPGSPTP- 239  
QY 239 TSFLPMGPPPAEGST-GDFALPVLGIVGTALGLLIIGVNCVIMTVQVKKPLCLQRE 297  
DB 240 -SILTSLGSTPIEQSTKGISLPIGLIVGTVSLGLLMLGLVNCIILVQKKPKPSCLQRD 298  
QY 298 AKYPHLPADKARGTOGPEQOHLITAPSSSSSSLESSASALDRRAPTRNQAPGV-EAS 356  
DB 299 AKYPHPVDEKSDQAVGLEQOHLITAPSSSSSSLESSASAGDRRAPGHPQARVMAEQ 358  
QY 357 GAGEARATGSSDSSPGHGTQVNTCIVNCSSSHSSQCSQASSTMGDTSSPSESP 416  
DB 359 GFQEARASSRISDSSHGSHGTHVNTCIVNCSSDSSSHSSQCSQASATVGDPAKPSASP 418  
QY 417 KDEQVPFSKEEACFRSQTETPTLLGSTTEKPLPLGVPDAGMKPS 461  
DB 419 KDEQVPFSQECPSQSPCETTTETL--QSHEKPLPLGVPDGMKPS 461

RESULT 11

US-08-385-229-4  
; Sequence 4, Application US/08385229  
; Patent No. 5605690  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Cindy A.  
; APPLICANT: Smith, Craig A.  
; TITLE OF INVENTION: Method of Treating TNF-Dependent  
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/385,229  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/946,236  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wright, Christopher L.  
; REGISTRATION NUMBER: 31,680  
; REFERENCE/DOCKET NUMBER: 2503  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 587-0606  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 518 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-385-229-4  
Query Match 58.1%; Score 1433; DB 1; Length 518;  
Best Local Similarity 64.5%; Pred. No. 2.5e-97;  
Matches 289; Conservative 22; Mismatches 67; Indels 70; Gaps 10;  
QY 1 MAPVAVMAALAVGLELWAAHALPAQVAFYAPPEGSCRLREYDDQTAQCCSKCSPG 60  
DB 30 MAPVAVMAALAVGLELWAAHALPAQVAFYAPPEGSCRLREYDDQTAQCCSKCSPG 89  
QY 61 QHAKVCTKTSYDVCDCEDSTYTQLMNWPCLSCGSCSDQVETQACTREQNRIC 120  
DB 90 QHAKVCTKTSYDVCDCEDSTYTQLMNWPCLSCGSCSDQVETQACTREQNRIC 149  
QY 121 RPHGYCALSKQBG-CRLCAPLRKCRPGFVARPGTETSDVCKPCAPGTFSTTSSTDICR 180  
DB 150 RPHGYCALSKQBG-CRLCAPLRKCRPGFVARPGTETSDVCKPCAPGTFSTTSSTDICR 209  
QY 181 PHQICNVVAIPGNASMDAVCTSTPTSRMAPCAVHLPOPVSTRSQHTOPTPEPSTAPSTS 240  
DB 210 PHQICNVVAIPGNASMDAVCTSTPTSRMAPCAVHLPOPVSTRSQHTOPTPEPSTAPSTS 269  
QY 241 FLLPMGSPPAEGSTGDFALPVLGIVGTALGLLIIGVNCVIMTVQVKKPKPLCLQREAKV 300  
DB 270 FLLPMGSPPAEGSTGDFALPVLGIVGTALGLLIIGVNCVIMTVQVKKPKPLCLQREAKV 300  
QY 301 PHLPADKARGTOGPEQOHLITAPSSSSSSLESSASALDRRAP-----TRNQOQ--- 349  
DB 298 PPCAPALLG--GPS----VFLFPKPKDITLMS-----RTPEVTCVVVDVSHEDPEVK 345  
QY 350 ----APGVEASGAGEARASTGSSDSSPGHGTQVNTCIVNV--CSSSDHSSQCSQA-S 402  
DB 346 FNMVYDVGVEVH-----NAKTKPREEQYNSTYRVVSVLVLHQLHODWLNKGDYCKKVSNNKALP 400  
QY 403 STMGDTSSPSESPKDEQV---PFSKEE 427  
DB 401 APMOKTISKAGQPREQVYTLPPSRDE 428  
RESULT 12  
US-08-243-010-1  
; Sequence 1, Application US/08243010  
; Patent No. 5639597  
; GENERAL INFORMATION:  
; APPLICANT: Laufer, Leander  
; APPLICANT: Zettlmeissel, Gerd  
; APPLICANT: Okendo, Patricia  
; TITLE OF INVENTION: Cell-free Receptor Binding Assays, The  
; TITLE OF INVENTION: Production and Use Thereof  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/243,010  
; FILING DATE: 13-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 07/798,564
; FILING DATE: 26-NOV-1991
; APPLICATION NUMBER: DE P 40 37 837.3
; FILING DATE: 28-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 02481-1132-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-243-010-1

Query Match          56.9%; Score 1404; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.1e-95;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPVAVMAALAVGLELWAAAHALPAQVAFTPYAPPEPGSTCRLREYYDQTQMCCSKCSPG 60
Db 1 MAPVAVMAALAVGLELWAAAHALPAQVAFTPYAPPEPGSTCRLREYYDQTQMCCSKCSPG 60
QY 61 QHAKVCTKTSDIVDSCEDSTYQLNNWVPECLSCGSRSSDQVETQACTREONRICTC 120
Db 61 QHAKVCTKTSDIVDSCEDSTYQLNNWVPECLSCGSRSSDQVETQACTREONRICTC 120
QY 121 RPYWCYCALSKOEGRCRLCAPLKRCPGFGVAPRGTTSDVVCKPCAPCTFSNTTSDTICR 180
Db 121 RPYWCYCALSKOEGRCRLCAPLKRCPGFGVAPRGTTSDVVCKPCAPCTFSNTTSDTICR 180
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPS 240
Db 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPS 240
QY 241 FILPMGSPPAE 252
Db 241 FILPMGSPPAE 252

RESULT 13
US-09-326-394-4
; Sequence 4, Application US/09326394
; Patent No. 6306820
; GENERAL INFORMATION:
; APPLICANT: Bendele, Alison M.
; APPLICANT: Sennello, Regina M.
; APPLICANT: Edwards, Carl K.
; TITLE OF INVENTION: COMBINATION THERAPY USING A TNF BINDING
; TITLE OF INVENTION: PROTEIN FOR TREATING TNF-MEDIATED DISEASES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: CA
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/326,394
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 60/032,587
; FILING DATE: 06-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,355
; FILING DATE: 23-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,315
; FILING DATE: 07-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/052,023
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Zindrick, Thomas K.
; REGISTRATION NUMBER: 32,185
; REFERENCE/DOCKET NUMBER: A-430D
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-326-394-4

Query Match          53.2%; Score 1312; DB 4; Length 235;
Best Local Similarity 99.6%; Pred. No. 6.8e-89;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 LPAQVAFTPYAPPEPGSTCRLREYYDQTQMCCSKCSPGQHAKVCTKTSDIVDSCEDST 82
Db 1 LPAQVAFTPYAPPEPGSTCRLREYYDQTQMCCSKCSPGQHAKVCTKTSDIVDSCEDST 60
QY 83 YTQLNNWVPECLSCGSRSSDQVETQACTREONRICTCRPGWYCALSKOEGRCRLCAPLKR 142
Db 61 YTQLNNWVPECLSCGSRSSDQVETQACTREONRICTCRPGWYCALSKOEGRCRLCAPLKR 120
QY 143 CRPGFGVAPRGTTSDVVCKPCAPCTFSNTTSDTICRPHQICNVVAIPGNASMDAVCTS 202
Db 121 CRPGFGVAPRGTTSDVVCKPCAPCTFSNTTSDTICRPHQICNVVAIPGNASMDAVCTS 180
QY 203 TSPTSRMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPSFLPMGSPPAEGSTGD 257
Db 181 TSPTSRMAPGAVHLPQVPVSTRSQHTQPTPEPSTAPSFLPMGSPPAEGSTGD 235

RESULT 14
US-08-974-022-48
; Sequence 48, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/577,788  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-378  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 227 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-974-022-48

Query Match 51.2%; Score 1263; DB 3; Length 227;  
Best Local Similarity 99.6%; Pred. No. 2.5e-85;  
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPGPGSTCRLREYYDQTAQMCCSKCSPG 60  
DB 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPGPGSTCRLREYYDQTAQMCCSKCSPG 60  
QY 61 QHAKVFCTKTSDFVCDSCDSTYTQLWNWVPECLSCGSRCSDDQVETQACTREQNRICTC 120  
DB 61 QHAKVFCTKTSDFVCDSCDSTYTQLWNWVPECLSCGSRCSDDQVETQACTREQNRICTC 120  
QY 121 RPYGWCALSKQEGCRCLCAPLRKCPGFGVARPGTETSDVVKPCAPGTFSTNTSSDIDICR 180  
DB 121 RPYGWCALSKQEGCRCLCAPLRKCPGFGVARPGTETSDVVKPCAPGTFSTNTSSDIDICR 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT 227  
DB 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT 227

RESULT 15

US-08-795-445A-48  
Sequence 48, Application US/08795445A  
Patent No. 6284485  
GENERAL INFORMATION:  
APPLICANT: Boyle, William J.  
APPLICANT: Lacey, David L.  
APPLICANT: Calzone, Frank J.  
APPLICANT: Chang, Ming-Shi  
TITLE OF INVENTION: OSTEOPROTEGERIN  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 Behavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,445A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/577,788  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-378  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 227 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-795-445A-48

Query Match 51.2%; Score 1263; DB 4; Length 227;  
Best Local Similarity 99.6%; Pred. No. 2.5e-85;  
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPGPGSTCRLREYYDQTAQMCCSKCSPG 60  
DB 1 MAPVAVMAALAVGLELWAAHALPAQVAFTPYAPGPGSTCRLREYYDQTAQMCCSKCSPG 60  
QY 61 QHAKVFCTKTSDFVCDSCDSTYTQLWNWVPECLSCGSRCSDDQVETQACTREQNRICTC 120  
DB 61 QHAKVFCTKTSDFVCDSCDSTYTQLWNWVPECLSCGSRCSDDQVETQACTREQNRICTC 120  
QY 121 RPYGWCALSKQEGCRCLCAPLRKCPGFGVARPGTETSDVVKPCAPGTFSTNTSSDIDICR 180  
DB 121 RPYGWCALSKQEGCRCLCAPLRKCPGFGVARPGTETSDVVKPCAPGTFSTNTSSDIDICR 180  
QY 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT 227  
DB 181 PHQICNVVAIPGNASMDAVCTSTSPTRSMAPGAVHLPQPVSTRSQHT 227

Search completed: February 12, 2002, 12:48:00  
Job time: 152 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 12, 2002, 13:02:20 ; Search time 12.79 Seconds  
(without alignments)  
136.983 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_185

Perfect score: 134

Sequence: 1 PCAPGTFSTSTSDICRPHQIC 23

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4466

Minimum DB seq length: 0  
Maximum DB seq length: 23

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	23.1	20	A42865	Ca2+/calmodulin-de
2	30	22.4	18	B49048	T-cell receptor be
3	29	21.6	13	S47381	T-cell antigen rec
4	29	21.6	16	S38292	30K allergen - rye
5	29	21.6	17	A60317	glucagon-like pept
6	29	21.6	21	I54351	gene HEXA protein
7	29	21.6	22	A39269	Lx-1 tumor antigen
8	28	20.9	23	S60565	homeodomain protei
9	27	20.1	20	T48881	leader peptide [im
10	27	20.1	22	MKN1	mu-conotoxin GIIIA
11	26	19.4	16	JN0263	antigen (clone PV1
12	26	19.4	15	B60278	24K antigen - Myco
13	26	19.4	20	I54283	arylsulfatase A -
14	26	19.4	23	A59048	convulsant peptide
15	25.5	19.0	22	I37144	aspartylglycosamin
16	25.5	19.0	23	E39855	paralytic peptide
17	25.5	19.0	23	D39855	paralytic peptide
18	25	18.7	11	S23308	substance P - rain
19	25	18.7	13	P00491	self-incompatibili
20	25	18.7	13	D56661	S-locus specific g
21	25	18.7	17	B61334	trypsin (EC 3.4.21
22	25	18.7	19	B56613	virion morphogenes
23	25	18.7	19	C56661	S-locus specific q
24	25	18.7	19	P00492	self-incompatibili
25	25	18.7	22	MYKN2	mu-conotoxin GIIIB
26	25	18.7	22	MYKN3	mu-conotoxin GIIIC
27	25	18.7	23	A48968	exo-poly-alpha-gal
28	25	18.7	23	B38671	peptidylglycine mo
29	24	17.9	11	PH0891	T-cell receptor be

30 24 17.9 14 2 PH0776 T-cell receptor al  
31 24 17.9 19 2 A05305 hemoglobin beta-2  
32 24 17.9 21 2 C39543 collagen alpha 3(I  
33 23.5 17.5 16 2 B54877 alpha-conotoxin Pn  
34 23.5 17.5 20 2 A34859 heloethermine - Mex  
35 23.5 17.5 23 2 F39855 paralytic peptide  
36 23.5 17.5 23 2 G39855 paralytic peptide  
37 23 17.2 12 1 JTJGO tremmerogen A-10 -  
38 23 17.2 12 2 I40663 bma protein - Cios  
39 23 17.2 13 2 G22365 R-phycoerythrin ga  
40 23 17.2 13 2 PS0453 36K protein 3124 -  
41 23 17.2 13 2 PH0799 T-cell receptor al  
42 23 17.2 14 2 B61309 lutropin beta chai  
43 23 17.2 15 2 A49155 vasotocin-associat  
44 23 17.2 15 2 PH0782 T-cell receptor al  
45 23 17.2 16 2 PH1634 Ig H chain V-D-J r

ALIGNMENTS

RESULT 1

A42865  
Ca2+/calmodulin-dependent myosin light chain kinase (autophosphorylation sites) - rab  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: A42865  
R:Gao, Z.H.; Moomaw, C.R.; Hsu, J.; Slaughter, C.A.; Stull, J.T.  
Biochemistry 31, 6126-6133, 1992  
A:Title: Autophosphorylation of skeletal muscle myosin light chain kinase.  
A:Reference number: A42865; MUID:92329432  
A:Accession: A42865  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <GAO>  
A:Experimental source: skeletal muscle  
A:Note: sequence extracted from NCBI backbone (NCBIP:109204)  
C:Keywords: calmodulin binding

Query Match 23.1% Score 31; DB 2; Length 20;  
Best Local Similarity 37.5%; Pred. No. 4.8e+02; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 9;  
QY 3 APGTFSTSTSDICR 18  
||| : |||  
Db 2 APGQADQAKAQGDTCR 17  
||| : |||  
RESULT 2  
B49048  
T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (Fragmen  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997  
C:Accession: B49048  
R:Sloud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.  
Eur. J. Immunol. 22, 2413-2418, 1992  
A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juven  
A:Reference number: A49048; MUID:92387250  
A:Accession: B49048  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-18 <SIO>  
A:Experimental source: patient EV, IL-2R+ synovial T-cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:113264)  
C:Keywords: T-cell receptor

Query Match 22.4% Score 30; DB 2; Length 18;  
Best Local Similarity 50.0%; Pred. No. 6e+02; Indels 0; Gaps 0;  
Matches 5; Conservative 1; Mismatches 4;  
QY 2 CAPGTFSTNTT 11

Db 7 CAPGXYYGT 16  
||||| : |

## RESULT 3

S47381

T-cell antigen receptor VJ junction beta chain - human

C:Species: Homo sapiens (man)

C&gt;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999

C:Accession: S47381

R:Lehner, P.J.

submitted to the EMBL Data Library, August 1994

A:Description: Human HLA-A\*0201 restricted recognition of influenza A is dominated by T

A:Reference number: S47355

A:Accession: S47381

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-13 &lt;LEH&gt;

A:Cross-references: EMBL:Z35698; NID:g527487; PIDN:CAA84767.1; PID:g527488

C:Keywords: T-cell receptor

## Query Match

Best Local Similarity 21.6%; Score 29; DB 2; Length 13;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CAPGTPSNT 10

||| | |||

Db 1 CASSRSNT 9

## RESULT 4

S38292

30K allergen - rye (fragment)

C:Species: Secale cereale (rye)

C&gt;Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 07-May-1999

C:Accession: S38292

R:Peterson, A.; Schramm, G.; Becker, W.M.; Schlaak, M.

Biol. Chem. Hoppe-Seyler 374, 855-861, 1993

A:Title: Comparison of four grass pollen species concerning their allergens of grass gro

A:Reference number: S38288; MUID:94092339

A:Accession: S38292

A:Molecule type: protein

A:Residues: 1-16 &lt;PET&gt;

## Query Match

Best Local Similarity 21.6%; Score 29; DB 2; Length 16;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PCAPGT 6

| |||||

Db 7 PAAPGT 12

## RESULT 5

A60317

glucagon-like peptide 1 - marbled electric ray (fragment)

C:Species: Torpedo marmorata (marbled electric ray)

C&gt;Date: 28-Oct-1992 #sequence\_revision 28-Oct-1992 #text\_change 21-Nov-1997

C:Accession: A60317

R:Conlon, J.M.; Hansen, H.F.; Schwartz, T.W.

Regul. Pept. 13, 94, 1986

A:Title: A truncated glucagon-like peptide I from torpedo pancreas.

A:Reference number: A60317

A:Accession: A60317

A:Molecule type: protein

A:Residues: 1-17 &lt;CON&gt;

C:Superfamily: glucagon

C:Keywords: duplication; pancreas

## Query Match

21.6%; Score 29; DB 2; Length 17;

Best Local Similarity 38.5%; Pred. No. 7.7e+02;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 APGTFSTSTSD 15

| ||| : || :

Db 2 AEGTYTSDVSLB 14

## RESULT 6

I54351

gene HEXA protein - human (fragment)

C:Species: Homo sapiens (man)

C&gt;Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jul-2000

C:Accession: I54351

R:Akli, S.; Chomel, J.C.; Lacorte, J.M.; Bachner, L.; Poenaru, A.; Poenaru, L.

Hum. Mol. Genet. 2, 61-67, 1993

A:Title: Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs patients.

A:Reference number: I54351; MUID:93258352

A:Accession: I54351

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-21 &lt;RES&gt;

A:Cross-references: GB:S61298; NID:g300412; PIDN:AAD13927.1; PID:g4261627

C:Genetics:

A:Gene: GDB:HEXA

A:Cross-references: GDB:I20040; OMIM:272800

A:Map position: I5q23-I5q24

C:Superfamily: beta-hexosaminidase

## Query Match

Best Local Similarity 21.6%; Score 29; DB 2; Length 21;

Matches 10; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy 5 GT-PSNTSTSDICR-PHQ 21

||| | | | | | | | |

Db 1 GTFFINKTEIEDFPFPHQ 19

## RESULT 7

A39269

LX-1 tumor antigen - human (fragment)

C:Species: Homo sapiens (man)

C&gt;Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 05-Jan-1996

C:Accession: A39269

R:Rosenbaum, L.C.; Newwelt, E.A.; Van Tol, H.H.M.; Loh, Y.P.; Verbalis, J.G.; Hellstr

Proc. Natl. Acad. Sci. U.S.A. 87, 9928-9932, 1990

A:Title: Expression of neurophysin-related precursor in cell membranes of a small-cel

A:Reference number: A39269; MUID:91088624

A:Accession: A39269

A:Molecule type: protein

A:Residues: 1-22 &lt;ROS&gt;

C:Superfamily: oxytocin-neurophysin

## Query Match

Best Local Similarity 21.6%; Score 29; DB 2; Length 22;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PCAPG 5

| | | |

Db 12 PCGPG 16

## RESULT 8

S60565

homeodomain protein hrox3 - California red abalone (fragment)

C:Species: Haliotis rufescens (California red abalone)

C&gt;Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 15-Oct-1999

C:Accession: S60565

R:Deegan, B.M.; Morse, D.E.

Mol. Marine Biol. Biotechnol. 2, 1-9, 1993

A:Title: Identification of eight homeobox-containing transcripts expressed during lar

A:Reference number: S60564; MUID:93372986  
 A:Accession: S60565  
 A:Molecule type: mRNA  
 A:Residues: 1-23 <DE>  
 A:Cross-references: EMBL:X79372; NID:9495110; PIDN:CAA55917.1; PID:9495111  
 C:Genetics:  
 A:Gene: hrox3  
 C:Superfamily: unassigned homeobox proteins; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 20.9%; Score 28; DB 2; Length 23;  
 Best Local Similarity 57.1%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 16 ICRPHQI 22  
 :||:|  
 Db 5 LCRPRRI 11

## RESULT 9

T48881  
 leader peptide [Imported] - Vibrio sp.  
 C:Species: Vibrio sp.  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 C:Accession: T48881  
 R:Xu, Y.; Zhang, Y.; Liang, Z.Y.; Van de Castele, M.; Legrain, C.; Glansdorff, N.  
 Microbiology 144, 1435-1441, 1998  
 A:Title: Aspartate carbamoyltransferase from a psychrophilic deep-sea bacterium, *Vibrio*  
 A:Reference number: 224845  
 A:Accession: T48881  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-20 <XUV>  
 A:Cross-references: EMBL:Y09786; PIDN:CAA70922.1  
 A:Experimental source: strain 2693

Query Match 20.1%; Score 27; DB 2; Length 20;  
 Best Local Similarity 41.7%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 9 NTSSTDICRPH 20  
 :||:|  
 Db 7 SSLSSFKLVRRH 18

## RESULT 10

MXKNI  
 mu-conotoxin GIIA [validated] - cone shell (*Conus geographus*)  
 N:Alternate names: geographotoxin I (GTx I); myotoxin I  
 C:Species: *Conus geographus* (geography cone)  
 C:Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 15-Sep-2000  
 C:Accession: A01786; A23579  
 R:Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.  
 FEBS Lett. 155, 277-280, 1983  
 A:Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins from  
 A:Reference number: A91309; MUID:83210170  
 A:Accession: A01786  
 A:Molecule type: protein  
 A:Residues: 1-22 <SAT>  
 R:Cruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczyld  
 J. Biol. Chem. 260, 9280-9288, 1985  
 A:Title: *Conus geographus* toxins that discriminate between neuronal and muscle sodium ch  
 A:Reference number: A23579; MUID:85261316  
 A:Accession: A23579  
 A:Molecule type: protein  
 A:Residues: 1-22 <CRU>  
 R:Kohda, D.; Lancelin, J.M.; Inagaki, F.; Wakamatsu, K.  
 submitted to the Brookhaven Protein Data Bank, December 1992  
 A:Reference number: A51994; PDB:1ITC  
 A:Contents: annotation; conformation by (1)H-NMR, residues 1-22  
 R:Lancelin, J.M.; Kohda, D.; Tate, S.I.; Yanagawa, Y.; Abe, T.; Satake, M.; Inagaki, F.

Biochemistry 30, 6908-6916, 1991  
 A:Title: Tertiary structure of conotoxin GIIIA in aqueous solution.  
 A:Reference number: A44659; MUID:91299744  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R:Ott, K.H.; Becker, S.; Gordon, R.D.; Rueterjans, H.  
 FEBS Lett. 278, 160-166, 1991  
 A:Title: Solution structure of mu-conotoxin GIIIA analysed by 2D-NMR and distance geo  
 A:Reference number: A58581; MUID:91122275  
 C:Contents: annotation; conformation by (1)H-NMR  
 R:Wakamatsu, K.; Kohda, D.; Hatanaka, H.; Lancelin, J.M.; Ishida, Y.; Oya, M.; Nakamu  
 Biochemistry 31, 12577-12584, 1992  
 A:Title: Structure-activity relationships of mu-conotoxin GIIIA: structure determinat  
 A:Reference number: A44244; MUID:93112598  
 C:Contents: annotation; conformation by (1)H-NMR  
 C:Superfamily: mu-conotoxin  
 C:Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor  
 F:3-15,4-20,10-21/disulfide bonds: #status experimental  
 F:6,7/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:17/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F:22/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 20.1%; Score 27; DB 1; Length 22;  
 Best Local Similarity 42.9%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 17 CRPHQIC 23  
 :||:|  
 Db 15 CKPQGCC 21

## RESULT 11

JN0263  
 antigen (clone PV12) - Plasmodium vivax (fragment)  
 C:Species: Plasmodium vivax  
 C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 09-Sep-1997  
 C:Accession: JN0263; S21344  
 R:Ray, P.; Sharma, Y.D.  
 Biochem. Biophys. Res. Commun. 184, 668-672, 1992  
 A:Title: Molecular cloning and serological characterization of a new Plasmodium vivax  
 A:Reference number: JN0263; MUID:92246949  
 A:Accession: JN0263  
 A:Molecule type: DNA  
 A:Residues: 1-15 <RAY>  
 A:Cross-references: GB:X53681; NID:gl0084; PID:gl0085

Query Match 19.4%; Score 26; DB 2; Length 15;  
 Best Local Similarity 54.5%; Pred. No. 1.8e+03;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PGTFSTNTSST 14  
 :||:|  
 Db 3 PGQHSETLVST 13

## RESULT 12

B60278  
 24K antigen - Mycobacterium bovis (fragment)  
 C:Species: Mycobacterium bovis  
 C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 18-Jun-1993  
 C:Accession: B60278  
 R:Pifis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.  
 Infect. Immun. 59, 800-807, 1991  
 A:Title: Purification and characterization of major antigens from a Mycobacterium bo  
 A:Reference number: A60278; MUID:91147217  
 A:Accession: B60278  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-16 <FIF>

Query Match 19.4%; Score 26; DB 2; Length 16;

Best Local Similarity 38.5%; Pred. No. 1.9e+03; Mismatches 7; Indels 0; Gaps 0;

QY 3 APTGFSNTTSSD 15  
|||  
Db 1 APTTYKEELKGT 13

## RESULT 13

I54283  
arylsulfatase A - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: I54283  
R:Regis. S.; Carozzo, R.; Filocamo, M.; Serra, G.; Mastropaolo, C.; Gatti, R.  
Hum. Genet. 96, 233-235, 1995  
A:Title: An AT-deletion causing a frameshift in the arylsulfatase A gene of a late infant  
A:Reference number: I54283; MUID:95362256  
A:Accession: I54283  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-20 <RES>  
A:Cross-references: GB:S78735; NID:gl037139; PIDN:AAB35013.1; PID:gl037140

Query Match 19.4%; Score 26; DB 2; Length 20;

Best Local Similarity 44.4%; Pred. No. 2.3e+03; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRP 19  
::: ||||  
Db 1 SAHSDHCRP 9

## RESULT 14

A59048  
convulsant peptide - cone shell (Conus textile)  
C:Species: Conus textile (cloth-of-gold cone)  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 13-Aug-1999  
C:Accession: A59048  
R:Cruz, L.J.; Ramilo, C.A.; Corpuz, G.P.; Olivera, B.M.  
Biol. Bull. 183, 159-164, 1992  
A:Title: Conus peptides: phylogenetic range of biological activity.  
A:Reference number: A59048  
A:Accession: A59048  
A:Molecule type: protein  
A:Residues: 1-23 <CRU>  
C:Keywords: amidated carboxyl end; neurotoxin; venom  
F:23/Modified site: amidated carboxyl end (pro) #status predicted

Query Match 19.4%; Score 26; DB 2; Length 23;

Best Local Similarity 27.8%; Pred. No. 2.6e+03; Mismatches 7; Indels 4; Gaps 1;

QY 2 CAPGFSNTTSTDICRP 19  
|||  
Db 9 CCPPAYCEASG---CRP 22

## RESULT 15

I37144  
aspartylglycosaminuria - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I37144  
R:Park, H.; Vettese, M.B.; Fensom, A.H.; Fisher, K.J.; Aronson, N.N.  
Biochem. J. 290, 735-741, 1993  
A:Title: Characterization of three alleles causing aspartylglycosaminuria: two from a Br  
A:Reference number: I37144; MUID:93207523  
A:Accession: I37144  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1-22 <RES>  
A:Cross-references: EMBL:X73071; NID:g312227; PIDN:CAA51529.1; PID:g312228  
C:Genetics:  
A:Gene: AGU

Query Match 19.0%; Score 25.5; DB 2; Length 22;  
Best Local Similarity 38.9%; Pred. No. 2.9e+03;  
Matches 7; Conservative 0; Mismatches 4; Indels 7; Gaps 1;

QY 6 TFSNTTSTDICRPHQIC 23  
||| |  
Db 5 TFSRRVS-----HHIC 15

Search completed: February 12, 2002, 13:04:11  
Job time: 111 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 13:03:56 ; Search time 10.06 Seconds  
(without alignments)  
83.826 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_185

Perfect score: 134

Sequence: 1 PCAPGTFSTSTSDICRPHQIC 23

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1274

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	20.1	22	1 CXM1_CONGE	P01523 Conus geogr
2	27	20.1	23	1 ACP2_BRANA	P30226 brassica na
3	25.5	19.0	23	1 CP23_SPOER	P56683 spodoptera
4	25.5	19.0	23	1 PAP2_SPOEX	P30256 spodoptera
5	25.5	19.0	23	1 PAP3_SPOEX	P30257 spodoptera
6	25	18.7	11	1 TKNA_ONCMY	P28499 oncorhynch
7	25	18.7	22	1 CXM2_CONGE	P01524 conus geogr
8	25	18.7	22	1 CXM3_CONGE	P05482 conus geogr
9	24.5	18.3	22	1 LANM_STRMU	P80666 streptococc
10	24	17.9	19	1 HBB2_UROHA	P19992 uromastix h
11	23.5	17.5	15	1 DCMW_PSECA	P19920 pseudomonas
12	23.5	17.5	16	1 CXAB_CONPE	P50985 conus penna
13	23.5	17.5	20	1 HELT_HELHO	P46693 helioderma h
14	23.5	17.5	23	1 PAP1_HELVI	P30251 heliothis v
15	23.5	17.5	23	1 PAP2_HELVI	P30252 heliothis v
16	23	17.2	12	1 TAI0_TREME	P01371 tremella me
17	23	17.2	15	1 CIOA_RAT	P31720 rattus norv
18	22.5	16.8	23	1 PAP2_MANSE	P30254 manduca sex
19	22	16.4	20	1 JHBP_BOMMO	P81627 bombyx mori
20	22	16.4	23	1 PAP1_SPOEX	P30255 spodoptera
21	21.5	16.0	23	1 CXAA_CONPE	P50984 conus penna
22	21	15.7	9	1 RS11_SALTY	O54296 salmonella
23	21	15.7	15	1 NUO3_SALTU	P80263 solanum tub
24	21	15.7	18	1 CXAL_CONER	P50982 conus ermin
25	21	15.7	20	1 PSBH_SYNVU	P19052 synechococc
26	21	15.7	22	1 SETB_SALTY	P33027 salmonella
27	21	15.7	22	1 TX12_TRIWA	P24335 trimeresuru
28	20	14.9	19	1 MIFH_TRISP	P81529 trichinella
29	20	14.9	20	1 UCRO_EQUAR	P81247 equisetum a
30	20	14.9	21	1 MCT3_MOUSE	P21843 mus musculu
31	20	14.9	23	1 RL5_HALHA	P50556 halobacteri
32	19.5	14.6	15	1 DCMW_PSECH	P19917 pseudomonas
33	19.5	14.6	23	1 PAP1_MANSE	P30253 manduca sex

#### ALIGNMENTS

##### RESULT 1

ID	CXML_CONGE	STANDARD;	PRT;	22 AA.
AC	P01523;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-FEB-1995 (Rel. 31, Last annotation update)			
DE	MU-CONOTOXIN GI1IA (MYOTOXIN I) (GEOGRAPHUTOXIN I) (GTX-I).			
OS	Conus geographus (Geography cone).			
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;			
OC	Neogastropoda; Conoidea; Conidae; Conus.			
OX	NCBI_TaxID=6491;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=85261316; PubMed=2410412;			
RA	Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,			
RA	Yoshikami D., Moczydlowski E.;			
RT	"Conus geographus toxins that discriminate between neuronal and			
RT	muscle sodium channels.";			
RL	J. Biol. Chem. 260:9280-9288(1985).			
RN	[2]			
RP	SEQUENCE.			
RX	MEDLINE=83210170; PubMed=6852238;			
RA	Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;			
RT	"The amino acid sequences of homologous hydroxyproline-containing			
RT	myotoxins from the marine snail Conus geographus venom.";			
RL	FEBS Lett. 155:277-280(1983).			
RN	[3]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=90249506; PubMed=2338142;			
RA	Hidaka Y., Sato K., Nakamura H., Kobayashi J., Ohizumi Y.,			
RA	Simonishi Y.;			
RT	"Disulfide pairings in geographutoxin I, a peptide neurotoxin from			
RT	Conus geographus.";			
RL	FEBS Lett. 264:29-32(1990).			
RN	[4]			
RP	REVIEW.			
RX	MEDLINE=89024586; PubMed=3052286;			
RA	Gray W.R., Olivera B.M., Cruz L.J.;			
RT	"Peptide toxins from venomous Conus snails.";			
RL	Annu. Rev. Biochem. 57:665-700(1988).			
RN	[5]			
RP	STRUCTURE BY NMR.			
RX	MEDLINE=9112275; PubMed=1991506;			
RA	Ott K.-H., Becker S., Gordon R.D., Rueterjans H.;			
RT	"Solution structure of mu-conotoxin GI1IA analysed by 2D-NMR and			
RT	distance geometry calculations.";			
RL	FEBS Lett. 278:160-166(1991).			
RN	[6]			
RP	STRUCTURE BY NMR.			
RX	MEDLINE=91299744; PubMed=2069951;			
RA	Lancelin J.-M., Kohda D., Tate S.-I., Yanagawa Y., Abe T., Satake M.,			
RA	Inagaki F.;			
RT	"Tertiary structure of conotoxin GI1IA in aqueous solution.";			
RL	Biochemistry 30:6308-6316(1991).			
CC	!- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK			

34	19	14.2	13	1	CXAL_CONST	P15471 conus stria
35	19	14.2	13	1	LMAL_LOCM1	P38496 locusta mig
36	19	14.2	15	1	CXAL_CONGE	P01519 conus geogr
37	19	14.2	15	1	HS11_PINPS	P81083 pinus pinas
38	19	14.2	15	1	NS11_ANASQ	Q44507 anabaena sp
39	19	14.2	15	1	RKGG_CARCR	P21586 caretta car
40	19	14.2	15	1	UC08_MAIZE	P80614 zea mays (m
41	19	14.2	15	1	URE2_MORMO	P17338 morganella
42	19	14.2	17	1	ITHE_HIRME	P28502 hirudo medi
43	19	14.2	19	1	CXA2_CONST	P28879 conus stria
44	19	14.2	19	1	LPRM_STAAN	P03063 staphylococ
45	19	14.2	20	1	SODE_PASPI	P81527 pasteurella

CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
CHANNELS.  
CC PIR; A01786; MXKN1.  
DR PIR; A23579; A23579.  
DR PDB; 1TCG; 31-JAN-94.  
DR PDB; 1TCH; 31-JAN-94.  
DR PDB; 1TCJ; 31-JAN-94.  
DR PDB; 1TCK; 31-JAN-94.  
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;  
3D-structure.  
FT DISULFID 3 15  
FT DISULFID 4 20  
FT DISULFID 10 21  
FT MOD.RES 6 6  
FT MOD.RES 7 7  
FT MOD.RES 17 17  
FT MOD.RES 22 22  
FT MOD.RES 13 16  
FT HELIX 19 21  
FT TURN 21  
SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;

Query Match 20.1%; Score 27; DB 1; Length 22;  
Best Local Similarity 42.9%; Pred. No. 5.1e+02;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 17 CRPHQIC 23  
|:|:|  
DB 15 CKPQRC 21

RESULT 2  
APP2\_BRANA STANDARD; PRT; 23 AA.  
AC P30226;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE CYSTEINE-RICH ANTIFUNGAL PROTEIN 2 (APP2) (FRAGMENT).  
OS Brassica napus (Rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3708;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Seed;  
RX MEDLINE=93138130; PubMed=8422949;  
RA Terras F.R.G., Torrekens S., van Leuven F., Osborn R.W.,  
RA Vanderleyden J., Cammue B.P.A., Broekaert W.F.;  
RT "A new family of basic cysteine-rich plant antifungal proteins from  
Brassicaceae species.";  
RL FEBS Lett. 316:233-240(1993).  
CC -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC  
CATIONS.  
CC -!- SUBUNIT: FORMS OLIGOMERS IN ITS NATIVE STATE.  
CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.  
DR PIR; S28992; S28992.  
DR HSP; P30231; IAYU.  
DR InterPro; IPR002118; Gamma-thionin.  
DR ProDom; PD002594; Gamma-thionin; 1.  
DR PROSITE; PS00940; GAMMA\_THIONIN; PARTIAL.  
KW Fungicide.  
FT MOD.RES 1 1  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2467 MW; A11D4A9E364F5735 CRC64;

Query Match 20.1%; Score 27; DB 1; Length 23;  
Best Local Similarity 28.6%; Pred. No. 5.3e+02;  
Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFQNTTSSTDICR 18

Db 9 GTWSGVCNNACK 22  
|:|:|

RESULT 3  
CP23\_SPOER STANDARD; PRT; 23 AA.  
AC P56683;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CARDIOACTIVE PEPTIDE CAP23.  
OS Spodoptera eridania (Southern armyworm).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.  
OX NCBI\_TaxID=37547;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=99196260; PubMed=10098624;  
RA Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,  
RA Shabanowitz J., Hunt D.F., Schooley D.A.;  
RT "A cardioactive peptide from the southern armyworm, Spodoptera  
eridania.";  
RL Peptides 20:53-61(1999).  
CC -!- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM  
LARVAL MANDUCA SEXTA. CAUSING AN INOTROPIC EFFECT AT LOW  
CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS  
AT HIGH DOSES.  
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.  
DR InterPro; IPR003463; GBP\_PSP.  
DR Pfam; PF02425; GBP\_PSP; 1.  
FT DISULFID 7  
FT SEQUENCE 23 AA; 2519 MW; 0A96D72A70855AE0 CRC64;  
BY SIMILARITY.

Query Match 19.0%; Score 25.5; DB 1; Length 23;  
Best Local Similarity 33.3%; Pred. No. 8.8e+02;  
Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPQTFNTTSSTDICRP 19  
|:|:|  
DB 7 CTPG-YQRTADGR--CKP 21

RESULT 4  
PAP2\_SPOEX STANDARD; PRT; 23 AA.  
AC P30256;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE PARALYTIC PEPTIDE II (PP II).  
OS Spodoptera exigua (Beet armyworm).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.  
OX NCBI\_TaxID=7107;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Hemolymph;  
RX MEDLINE=91302298; PubMed=2071576;  
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,  
RA Quistad G.B.;  
RT "Isolation and identification of paralytic peptides from hemolymph of  
the lepidopteran insects Manduca sexta, Spodoptera exigua, and  
Heliothis virescens.";  
RL J. Biol. Chem. 266:12873-12877(1991).  
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO  
LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE  
HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.  
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.  
DR PIR; D39855; D39855.

DR InterPro: IPR003463; GBP\_PSP.  
 DR Pfam: PF02425; GBP\_PSP; 1.  
 KW Hemolymph.  
 FT DISULFID 7 19  
 SQ SEQUENCE 23 AA; 2477 MW; 0A96CB4600855AE0 CRC64;

BY SIMILARITY.

Query Match 19.0%; Score 25.5; DB 1; Length 23;  
 Best Local Similarity 33.3%; Pred. No. 8.8e+02;  
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

OY 2 CAPGTFSTSTSDICRP 19  
 I I I : I : I I  
 Db 7 CTPG-YQTADGR--CKP 21

# RESULT 5

ID P3P3\_SPOEX STANDARD; PRT; 23 AA.  
 AC P30257;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE PARALYTIC PEPTIDE III (PP III).  
 OS Spodoptera exigua (Beet armyworm).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 CC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.  
 OX NCBI\_TaxID=7107;  
 RN [1]  
 RP SEQUENCE.  
 RC Tissue=Hemolymph;  
 RX MEDLINE=91302298; PubMed=2071576;  
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,  
 RA Quistad G.B.;  
 RT "Isolation and identification of paralytic peptides from hemolymph of  
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and  
 RT Heliothis virescens."  
 RL J. Biol. Chem. 266:12873-12877(1991).  
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO  
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE  
 CC HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING. FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.  
 DR PIR: E39855; E39855.  
 DR InterPro: IPR003463; GBP\_PSP.  
 DR Pfam: PF02425; GBP\_PSP; 1.  
 KW Hemolymph.  
 FT DISULFID 7 19  
 SQ SEQUENCE 23 AA; 2505 MW; 0A96CB5EB7D55AE0 CRC64;

Query Match 19.0%; Score 25.5; DB 1; Length 23;  
 Best Local Similarity 33.3%; Pred. No. 8.8e+02;  
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

OY 2 CAPGTFSTSTSDICRP 19  
 I I I : I : I I  
 Db 7 CTPG-YQTADGR--CKP 21

# RESULT 6

ID TKNA\_ONCMY STANDARD; PRT; 11 AA.  
 AC P28499;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;

# RN [1]

RP Tissue=Brain;  
 RC MEDLINE=92298992; PubMed=1376687;  
 RX Jensen J., Conlon J.M.;  
 RA "Substance-P-related and neurokinin-A-related peptides from the brain  
 RT of the cod and trout."  
 RL Eur. J. Biochem. 206:659-664(1992).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR: S23307; S23307.  
 DR PIR: S23308; S23308.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 KW MOD\_RES 11 11 AMIDATION (BY SIMILARITY).  
 FT SEQUENCE 11 AA; 1358 MW; 214860DEC9D6D1F7 CRC64;  
 SQ

Query Match 18.7%; Score 25; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 RPHQ 21  
 I I I I  
 Db 3 RPHQ 6

# RESULT 7

ID CXM2\_CONGE STANDARD; PRT; 22 AA.  
 AC P01524;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MU-CONOTOXIN GLIIB (MYOTOXIN II) (GEOGRAPHUTOXIN II) (GTX-II).  
 OS Conus geographus (Geography cone).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 CC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=6491;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=85261316; PubMed=2410412;  
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,  
 RA Yoshikami D., Moczydlowski E.;  
 RT "Conus geographus toxins that discriminate between neuronal and  
 RT muscle sodium channels."  
 RL J. Biol. Chem. 260:9280-9288(1985).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=83210170; PubMed=6852238;  
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;  
 RT "The amino acid sequences of homologous hydroxyproline-containing  
 RT myotoxins from the marine snail Conus geographus venom."  
 RL FEBS Lett. 155:277-280(1983).  
 RN [3]  
 RP REVIEW.  
 RX MEDLINE=89024586; PubMed=3052286;  
 RA Gray W.R., Olivera B.M., Cruz L.J.;  
 RT "Peptide toxins from venomous Conus snails."  
 RL Annu. Rev. Biochem. 57:665-700(1988).  
 RN [4]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=96280640; PubMed=8688418;  
 RA Hill J.M., Alewood P.F., Craik D.J.;  
 RT "Three-dimensional solution structure of mu-conotoxin GIIB, a  
 RT specific blocker of skeletal muscle sodium channels.";

RL Biochemistry 35:8824-8835(1996).  
CC -I- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
CC CHANNELS.

DR PIR: A01787; MXKN2  
DR PIR: B23579; B23579.  
DR PDB: 1GIB; 08-NOV-96.  
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;  
KW 3D-structure.  
FT DISULFID 3 15  
FT DISULFID 4 20  
FT DISULFID 10 21  
FT MOD\_RES 6 6 HYDROXYLATION.  
FT MOD\_RES 7 7 HYDROXYLATION.  
FT MOD\_RES 17 17 HYDROXYLATION.  
FT MOD\_RES 22 22 AMIDATION.  
SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;

Query Match 18.7%; Score 25; DB 1; Length 22;  
Best Local Similarity 44.4%; Pred. No. 9.9e+02;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 DICRPHQIC 23  
| | | | |  
Db 2 DCCTPPKC 10

RESULT 8  
CCX3\_CONGE STANDARD; PRT; 22 AA.  
AC P05482;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-AUG-1990 (Rel. 15, Last annotation update)  
DE MU-CONOTOXIN GIIC.  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6491;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=85261316; PubMed=2410412;  
RA Cruz L.J., Gray W.R., Olivera B.M., Zelkus R.D., Kerr L.,  
RA Yoshikami D., Moczydlowski E.;  
RT "Conus geographus toxins that discriminate between neuronal and  
RT muscle sodium channels.";  
RL J. Biol. Chem. 260:9280-9288(1985).  
RN [2]  
RP REVIEW.  
RX MEDLINE=89024586; PubMed=3052286;  
RA Gray W.R., Olivera B.M., Cruz L.J.;  
RT "Peptide toxins from venomous Conus snails.";  
RL Annu. Rev. Biochem. 57:665-700(1988).  
CC -I- FUNCTION: MU-CONOTOXINS ACT ON MUSCLE MEMBRANES. THEY BLOCK  
CC MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE-ACTIVATED SODIUM  
CC CHANNELS.

DR PIR: C23579; C23579.  
DR HSP: P01524; 1GIB.  
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom.  
FT DISULFID 3 15 BY SIMILARITY.  
FT DISULFID 4 20 BY SIMILARITY.  
FT DISULFID 10 21 BY SIMILARITY.  
FT MOD\_RES 6 6 HYDROXYLATION.  
FT MOD\_RES 7 7 HYDROXYLATION.  
FT MOD\_RES 17 17 HYDROXYLATION.  
FT MOD\_RES 22 22 AMIDATION.  
SQ SEQUENCE 22 AA; 2553 MW; F50402BA92A9813C CRC64;

Query Match 18.7%; Score 25; DB 1; Length 22;  
Best Local Similarity 44.4%; Pred. No. 9.9e+02;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 DICRPHQIC 23  
| | | | |  
Db 2 DCCTPPKC 10

RESULT 9  
LANM\_STRMU STANDARD; PRT; 22 AA.  
AC P80666;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LANTIBIOTIC MUTACIN B-NY266.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN [1]  
RP SEQUENCE.  
RX STRAIN=NY266;  
RX MEDLINE=97379322; PubMed=9237644;  
RA Mota-Meira M., Lacroix C., Lapointe G., Lavoie M.C.;  
RT "Purification and structure of mutacin B-NY266: a new lantibiotic  
RT produced by Streptococcus mutans.";  
RL FEBS Lett. 410:275-279(1997).  
CC -I- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)  
CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF  
CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL  
CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS  
CC TRANSMEMBRANE PORES.  
CC -I- MASS SPECTROMETRY: MW=2270.29; MW\_ERR=0.21; METHOD=ELECTROSPRAY.  
CC -I- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.  
DR InterPro: IPR001049; Gallidermin.  
DR Pfam: PF02052; Gallidermin; 1.  
DR PRINTS: PR00323; GALLIDERMIN.  
KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.  
FT MOD\_RES 5 5 DHA (2,3-DIDEHYDROALANINE).  
FT MOD\_RES 8 8 D-ABU (AMINOBUITYRIC ACID).  
FT MOD\_RES 14 14 DHB (2,3-DIDEHYDROBUTYRINE).  
FT MOD\_RES 19 19 DHA (2,3-DIDEHYDROALANINE).  
FT THIOETH 3 7 ALA-S-ALA (LANTHIONINE).  
FT THIOETH 8 11 ABU-S-ALA (BETA-METHYLLANTHIONINE).  
FT THIOETH 16 21 ALA-S-ALA (LANTHIONINE).  
FT THIOETH 19 22 DHA-S-ALA (AVI).  
SQ SEQUENCE 22 AA; 2425 MW; 961C1480401F92CE CRC64;

Query Match 18.3%; Score 24.5; DB 1; Length 22;  
Best Local Similarity 37.5%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 2 CAPGTFSTNTSSSTDIC 17  
| | | | |  
Db 7 CTGCG-AKTGSFNSYC 21

RESULT 10  
HBB2\_UROHA STANDARD; PRT; 19 AA.  
AC P18992;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-MAY-1991 (Rel. 18, Last annotation update)  
DE HEMOGLOBIN BETA-2 CHAIN (FRAGMENT).  
OS Uromastix hardwickii (Indian spiny-tailed lizard).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Lepidosaurs; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;  
OC Uromastix.  
OX NCBI\_TaxID=40250;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=84029159; PubMed=6628672;

RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,  
 RA Joernvall H.;  
 RT "Characterization of hemoglobin from the lizard Uromastix  
 RT hardwickii.";  
 RL FEBS Lett. 162:290-295(1983).  
 DR PIR; A05305; A05305.  
 DR InterPro: IPR000971; Globin.  
 DR PROSITE: PS01033; GLOBIN; PARTIAL.  
 KW Heme; Oxygen transport; Respiratory protein; Erythrocyte.  
 FT NON\_TER 1 1  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 17.9%; Score 24; DB 1; Length 19;  
 Best Local Similarity 41.7%; Pred. No. 1.2e+03;  
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 GTFSNTSSTDI 16  
 Db 3 GDFGNISAAAI 14  
 |||:|:|  
 |||:|:|

RESULT 11  
 DCMW\_PSECA STANDARD; PRT; 15 AA.  
 AC P19920;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] MEDIUM CHAIN (EC 1.2.2.4)  
 DE (FRAGMENT).  
 OS Pseudomonas carboxydovorans.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Bradyrhizobium group; Oligotropa.  
 OX NCBI\_TaxID=40137;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN-OMS;  
 RX MEDLINE=90055678; PubMed=2818128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 RT carboxydotropic bacteria.";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -!- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 +  
 CC 2 H(+) + FERRICYTOCHROME B-561.  
 CC -!- COFACTOR: MOLYBDENUM.  
 CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 CC SMALL.  
 DR PIR; P10141; P10141.  
 KW Oxidoreductase; Molybdenum.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1779 MW; 82DD3BF93E739D63 CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 15;  
 Best Local Similarity 36.8%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 4 PGTFNTSSTDI CRPHQI 22  
 ||:| | | |  
 Db 4 PGSF-----DYHRPKSI 15

RESULT 12  
 CXAB\_CONPE STANDARD; PRT; 16 AA.  
 ID CXAB\_CONPE  
 AC P50985;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE ALPHA-CONOTOXIN PNIB.  
 OS Conus pennaceus.

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=37335;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Venom;  
 MEDLINE=94347719; PubMed=8068627;  
 RA Fainzilber M., Hasson A., Oren R., Burlingame A.L., Gordon D.,  
 RA Fainzilber M., Zlotkin E.;  
 RT "New mollusc-specific alpha-conotoxins block Aplysia neuronal  
 RT acetylcholine receptors.";  
 RL Biochemistry 33:9523-9529(1994).  
 RN [2]  
 RP SULFATION OF TYR-15.  
 RX MEDLINE=99242956; PubMed=10226369;  
 RA Wolfender J.L., Chu F., Ball H., Wolfender F., Fainzilber M.,  
 RA Baldwin M.A., Burlingame A.L.;  
 RT "Identification of tyrosine sulfation in Conus pennaceus conotoxins  
 RT alpha-PnIA and alpha-PnIB: further investigation of labile sulfo- and  
 RT phosphopeptides by electrospray, matrix-assisted laser  
 RT desorption/ionization (MALDI) and atmospheric pressure MALDI mass  
 RT spectrometry.";  
 RL J. Mass Spectrom. 34:447-454(1999).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
 RX MEDLINE=97444322; PubMed=9298951;  
 RA Hu S.H., Gehrman J., Alewood P.F., Craik D.J., Martin J.L.;  
 RT "Crystal structure at 1.1-A resolution of alpha-conotoxin PnIB:  
 RT comparison with alpha-conotoxins PnIA and GI.";  
 RL Biochemistry 36:11323-11330(1997).  
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
 CC INHIBIT THEM. IN CONTRAST TO OTHER ALPHA-CONOTOXINS, WHICH ARE  
 CC SELECTIVE FOR VERTEBRATE SKELETAL MUSCLE NACHR, THE CONUS  
 CC PENNACEUS ALPHA-CONOTOXINS BLOCK NACHR IN MOLLUSCS.  
 DR PDB; LAG; 20-MAY-98.  
 KW postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
 KW Sulfation; Venom; 3D-structure.  
 FT DISULFID 2 8  
 FT DISULFID 3 16  
 FT MOD\_RES 15 15  
 FT MOD\_RES 16 16  
 SQ SEQUENCE 16 AA; 1643 MW; 05310FF95ED86AF5 CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 16;  
 Best Local Similarity 35.3%; Pred. No. 1.2e+03;  
 Matches 6; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 1 PCAFGTFSTNTSSTDIC 17  
 ||| | | | |  
 Db 7 PCA-----LSNPDYC 16

RESULT 13  
 HELT\_HELHO STANDARD; PRT; 20 AA.  
 ID HELT\_HELHO  
 AC P46693;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE HELOTHERMINE (FRAGMENT).  
 OS Heloderma horridum horridum (Mexican beaded lizard).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Anguilliformia; Helodermatidae;  
 OC Heloderma.  
 OX NCBI\_TaxID=8552;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Venom;  
 MEDLINE=90260878; PubMed=1693019;  
 RA Mochca-Morales J., Martin B.M., Possani L.D.;  
 RT "Isolation and characterization of helothermine, a novel toxin from

RT Heloderma horridum horridum (Mexican beaded lizard) venom."  
 RL Toxinon 28:299-309(1990).  
 CC -1- FUNCTION: TOXIC TO MICE; INDUCES LETHARGY, PARTIAL PARALYSIS OF  
 CC REAR LIMBS AND LOWERING OF BODY TEMPERATURE, SUGGESTING THAT IT  
 CC MIGHT BE A HYPOTHERMIC TOXIN.  
 CC -1- MISCELLANEOUS: THE COMPLETE PROTEIN HAS AN APPARENT MW OF 25 KDA  
 CC AND A PI OF 6.8.  
 DR PIR; A34859; A34859.  
 KW Toxin.  
 FT NON\_TER  
 SQ SEQUENCE 20 AA; 2156 MW; 91D62B36F7B4F940 CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 20;  
 Best Local Similarity 41.2%; Pred. No. 1.5e+03;  
 Matches 7; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 1 PCAPGTF-SNVTSTDI 16  
 I I I I I I  
 Db 4 PKLPLGLMTSNPDQQTET 20

RESULT 14  
 PAPI\_HELVI STANDARD; PRT; 23 AA.  
 AC P30251;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE PARALYTIC PEPTIDE I (PP I).  
 OS Heliothis virescens (Noctuid moth) (Owlet moth).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 CC Noctuoidea; Noctuidae; Heliothinae; Heliothis.  
 OX NCBI\_TaxID=7102;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Hemolymph;  
 RX MEDLINE=91302298; PubMed=2071576;  
 RA Skinner W.S.; Dennis P.A.; Li J.P.; Summerfelt R.M.; Carney R.L.;  
 RA Quistad G.B.;  
 RT "Isolation and identification of paralytic peptides from hemolymph of  
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and  
 RT Heliothis virescens."  
 RL J. Biol. Chem. 266:12873-12877(1991).  
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO  
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE  
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.  
 CC PIR; F39855; F39855.  
 DR InterPro: IPR003463; GBP\_PSP.  
 DR Pfam: PF02425; GBP\_PSP; I.  
 KW Hemolymph.  
 FT DISULFID 7 19 BY SIMILARITY.  
 SQ SEQUENCE 23 AA; 2524 MW; 2236CB436D655AFA CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 23;  
 Best Local Similarity 33.3%; Pred. No. 1.7e+03;  
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPGTFSTSTSDICRP 19  
 I I I I I I  
 Db 7 CIPG-YMRTADGR--CKP 21

RESULT 15  
 PAP2\_HELVI STANDARD; PRT; 23 AA.  
 AC P30252;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PARALYTIC PEPTIDE II (PP II).  
 OS Heliothis virescens (Noctuid moth) (Owlet moth).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 CC Noctuoidea; Noctuidae; Heliothinae; Heliothis.  
 OX NCBI\_TaxID=7102;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Hemolymph;  
 RX MEDLINE=91302298; PubMed=2071576;  
 RA Skinner W.S.; Dennis P.A.; Li J.P.; Summerfelt R.M.; Carney R.L.;  
 RA Quistad G.B.;  
 RT "Isolation and identification of paralytic peptides from hemolymph of  
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and  
 RT Heliothis virescens."  
 RL J. Biol. Chem. 266:12873-12877(1991).  
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO  
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE  
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.  
 CC PIR; G39855; G39855.  
 DR InterPro: IPR003463; GBP\_PSP.  
 DR Pfam: PF02425; GBP\_PSP; I.  
 KW Hemolymph.  
 SQ SEQUENCE 23 AA; 2508 MW; 2236CB5D6C855AFA CRC64;

Query Match 17.5%; Score 23.5; DB 1; Length 23;  
 Best Local Similarity 33.3%; Pred. No. 1.7e+03;  
 Matches 6; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 2 CAPGTFSTSTSDICRP 19  
 I I I I I I  
 Db 7 CIPG-YMRTADGR--CKP 21

Search completed: February 12, 2002, 13:06:44  
 Job time: 168 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 13:04:16 ; Search time 22.01 seconds  
(without alignments)  
152.852 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_185

Perfect score: 134

Sequence: 1 PCAPGTFSTSTSDICRPHQIC 23

Scoring table: BLOSUM62

Gapop 10.0., Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 6766

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP\_TREMBL.17.\*  
2: SP\_Archea.\*  
3: SP\_Bacteria.\*  
4: SP\_Fungi.\*  
5: SP\_Human.\*  
6: SP\_Invertebrate.\*  
7: SP\_MHC.\*  
8: SP\_Organelle.\*  
9: SP\_Phage.\*  
10: SP\_Plant.\*  
11: SP\_Rodent.\*  
12: SP\_Virus.\*  
13: SP\_Vertebrate.\*  
14: SP\_Unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	24.6	15	11 Q9QV5	Q9QV5 rattus sp.
2	30	22.4	20	12 Q78505	Q78505 human immun
3	29	21.6	17	12 Q78381	Q78381 human immun
4	29	21.6	18	12 Q9QEX3	Q9QEX3 human immun
5	29	21.6	21	4 Q16017	Q16017 homo sapien
6	28	20.9	17	12 Q78323	Q78323 human immun
7	28	20.9	17	12 Q78378	Q78378 human immun
8	28	20.9	17	12 Q78327	Q78327 human immun
9	28	20.9	17	12 Q78380	Q78380 human immun
10	28	20.9	21	5 Q25086	Q25086 herdmania m
11	28	20.9	23	5 Q25134	Q25134 haliotis ru
12	27	20.1	17	12 Q78379	Q78379 human immun
13	27	20.1	20	2 P96173	P96173 vibrio sp.
14	27	20.1	20	5 Q9TWR5	Q9TWR5 phoneutria
15	27	20.1	22	13 Q91102	Q91102 morone saxa
16	27	20.1	23	12 Q86611	Q86611 human immun
17	26	19.4	15	5 Q26159	Q26159 plasmodium
18	26	19.4	20	6 Q9TRH7	Q9TRH7 canis faml.
19	26	19.4	20	10 Q9S885	Q9S885 lupinus alb

20	26	19.4	20	11	Q9QV5	Q9QV5 rattus sp.
21	26	19.4	21	3	Q9UTT6	Q9UTT6 schizosacch
22	26	19.4	22	3	Q9UR51	Q9UR51 filobasidi
23	26	19.4	23	5	P90716	P90716 berce ovata
24	26	19.4	23	12	Q9QEX5	Q9QEX5 human immun
25	26	19.4	23	12	Q9QEX4	Q9QEX4 human immun
26	25.5	19.0	22	4	Q13726	Q13726 homo sapien
27	25	18.7	13	10	Q9S922	Q9S922 brassica ol
28	25	18.7	13	11	P97944	P97944 mus musculus
29	25	18.7	17	5	Q9TWC6	Q9TWC6 diroflaria
30	25	18.7	18	4	Q9UCN8	Q9UCN8 homo sapien
31	25	18.7	19	10	Q9S923	Q9S923 brassica ol
32	25	18.7	20	12	Q78486	Q78486 human immun
33	25	18.7	20	12	Q78507	Q78507 human immun
34	25	18.7	20	12	Q78508	Q78508 human immun
35	25	18.7	20	12	Q78509	Q78509 human immun
36	25	18.7	22	3	Q9Y8F8	Q9Y8F8 glomus moss
37	24.5	18.3	15	4	Q93046	Q93046 homo sapien
38	24	17.9	15	2	O69142	O69142 streptococc
39	24	17.9	15	10	Q9S8B9	Q9S8B9 lupinus alb
40	24	17.9	16	4	Q9UC48	Q9UC48 homo sapien
41	24	17.9	16	4	Q9NPO7	Q9NPO7 homo sapien
42	24	17.9	16	13	Q9PRU6	Q9PRU6 gallus gall
43	24	17.9	17	12	O85719	O85719 reovirus sp
44	24	17.9	17	12	Q78345	Q78345 human immun
45	24	17.9	17	13	Q9PRU7	Q9PRU7 gallus gall

## ALIGNMENTS

RESULT 1

Q9QV5 ID Q9QV5 PRELIMINARY; PRT; 15 AA.

AC Q9QV5; DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE: CERAMIDE GALACTOSYLTRANSFERASE

DE (FRAGMENT).

OS Rattus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10118;

RN [1]

RP MEDLINE=96085162; PubMed=8521863;

RA Schulte S., Stoffel W.

RT "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate

RT transporter. Copurification, separation and characterization of the

RT two glycoproteins."

RL Eur. J. Biochem. 233:947-953(1995).

SQ SEQUENCE 15 AA; 1657 MW; 84474749A06BF6CC CRC64;

Query Match 24.6%; Score 33; DB 11; Length 15;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PGTFSTNTS 12

||| :|||

Db 2 PGIFXSTTS 10

RESULT 2

Q78505 ID Q78505 PRELIMINARY; PRT; 20 AA.

AC Q78505; DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE VIRAL SAMPLE FLQ5R5D (FLORIDA LOCAL CONTROL 01), PARTIAL ENV CDS, V5

DE REGION (FRAGMENT).

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
Econoum A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92150; AAA44592.1; -.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2049 MW; F44F963A48755A07 CRC64;

Query Match 22.4%; Score 30; DB 12; Length 20;
Best Local Similarity 40.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSNTTSTTDICRP 19
Db 2 GNKNGTENETEIRFP 16

RESULT 3
Q78381 PRELIMINARY; PRT; 17 AA.
AC Q78381;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE VIRAL SAMPLE FLUPBR5F (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION
DE (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang L.Q., Leigh-Brown A.J.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
MEDLINE=92271245; PubMed=1589796;
RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,
Econoum A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,
Curran J.W., Jaffe H.W.;
RT "Molecular epidemiology of HIV transmission in a dental practice.";
RL Science 256:1165-1171(1992).
DR EMBL; M92126; AAA44496.1; -.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1708 MW; 347570D2D12CA370 CRC64;

Query Match 21.6%; Score 29; DB 12; Length 17;
Best Local Similarity 45.5%; Pred. No. 6.7e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 NNTSSTTDICRP 19
Db 3 NNTNNTETFRP 13

RESULT 4
Q9QEX3 PRELIMINARY; PRT; 17 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)

```

```

ID Q9QEX3 PRELIMINARY; PRT; 18 AA.
AC Q9QEX3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin H.J., Siwak E.B., Hollinger F.B.;
RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA
deduced from long term culture of its biological clones.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF178667; AAF04373.1; -.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 2011 MW; E17BAC9DD31D9910 CRC64;

Query Match 21.6%; Score 29; DB 12; Length 18;
Best Local Similarity 42.9%; Pred. No. 7.1e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 CAPGTFSTSTSTD 15
Db 2 CTCLNVTNTTISTE 15

RESULT 5
Q16017 PRELIMINARY; PRT; 21 AA.
AC Q16017;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE HEXA PROTEIN (FRAGMENT).
GN HEXA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93258352; PubMed=8490625;
RA Akli S., Chomel J.C., Lacorte J.M., Bachner L., Poenaru A.,
Poenaru L.;
RT "Ten novel mutations in the HEXA gene in non-Jewish Tay-Sachs
patients.";
RL Hum. Mol. Genet. 2:61-67(1993).
DR EMBL; S61298; AAD13927.1; -.
DR HSSP; P06865; IQBC.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2494 MW; D4ACE2D1DA24D8EC CRC64;

Query Match 21.6%; Score 29; DB 4; Length 21;
Best Local Similarity 52.6%; Pred. No. 8.2e+02;
Matches 10; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 5 GT-FSNTTSTTDICR-PHQ 21
Db 1 GTFFINKTEIDFPRPHQ 19

RESULT 6
Q78323 PRELIMINARY; PRT; 17 AA.
ID Q78323
AC Q78323;
DT 01-NOV-1996 (TREMBlrel. 01, Created)

```

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLPAR5A (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92271245; PubMed=1589796;  
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,  
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,  
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,  
 RA Curran J.W., Jaffe H.W.;  
 RT "Molecular epidemiology of HIV transmission in a dental practice."  
 RL Science 256:1165-1171(1992).  
 DR EMBL; M92109; AAA44465.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1649 MW; 3B857BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 12; Length 17;  
 Best Local Similarity 45.5%; Pred. No. 9.6e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19  
 I I I I I  
 Db 3 NNTNGTETFRP 13

RESULT 7  
 Q78378  
 ID Q78378 PRELIMINARY; PRT; 17 AA.  
 AC Q78378;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLPAR5A (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92271245; PubMed=1589796;  
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,  
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,  
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,  
 RA Curran J.W., Jaffe H.W.;  
 RT "Molecular epidemiology of HIV transmission in a dental practice."  
 RL Science 256:1165-1171(1992).  
 DR EMBL; M92123; AAA44493.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1723 MW; 34757935D12CA370 CRC64;

Query Match 20.9%; Score 28; DB 12; Length 17;  
 Best Local Similarity 45.5%; Pred. No. 9.6e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19  
 I I I I I  
 Db 3 NNTNGTETFRP 13

RESULT 8  
 Q78327  
 ID Q78327 PRELIMINARY; PRT; 17 AA.  
 AC Q78327;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLPAR5D (FLORIDA PATIENT A), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92271245; PubMed=1589796;  
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,  
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,  
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,  
 RA Curran J.W., Jaffe H.W.;  
 RT "Molecular epidemiology of HIV transmission in a dental practice."  
 RL Science 256:1165-1171(1992).  
 DR EMBL; M92112; AAA44468.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1651 MW; 34757BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 12; Length 17;  
 Best Local Similarity 45.5%; Pred. No. 9.6e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTSSTDICRP 19  
 I I I I I  
 Db 3 NNTNGTETFRP 13

RESULT 9  
 Q78380  
 ID Q78380 PRELIMINARY; PRT; 17 AA.  
 AC Q78380;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE VIRAL SAMPLE FLPAR5E (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION  
 DE (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang L.Q., Leigh-Brown A.J.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92271245; PubMed=1589796;  
 RA Ou C.Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,  
 RA Korber B.T.M., Mullins J.I., Schochetman G., Berkman R.L.,  
 RA Economou A.N., Witte J.J., Furman L.J., Satten G.A., MacInnes K.A.,  
 RA Curran J.W., Jaffe H.W.;  
 RT "Molecular epidemiology of HIV transmission in a dental practice."  
 RL Science 256:1165-1171(1992).  
 DR EMBL; M92125; AAA44495.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1651 MW; 34757BBFD12CA370 CRC64;

Query Match 20.9%; Score 28; DB 12; Length 17;  
 Best Local Similarity 45.5%; Pred. No. 9.6e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Query Match 20.9%; Score 28; DB 12; Length 17;  
Best Local Similarity 45.5%; Pred. No. 9.6e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NNTSSTDICRP 19  
| | | | |  
Db 3 NNTNGTETFRP 13

RESULT 10

Q25086 PRELIMINARY; PRT; 21 AA.

AC Q25086;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE CLONE AHOX4 HOMEBOX PROTEIN (FRAGMENT).  
OS Herdmania momus.  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
OC Stolidobranchia; Pyuridae; Herdmania.  
OX NCBI\_TaxID=7733;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-EMBRYO;  
RA Kennett C.V.D.;  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U09939; AAA18629.1; -.  
DR HSSP; P02833; 9ANT.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; Homeobox; 1.  
KW Homeobox; DNA-binding; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2650 MW; AB7FF3AF1FA659C3 CRC64;

Query Match 20.9%; Score 28; DB 5; Length 21;  
Best Local Similarity 57.1%; Pred. No. 1.2e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 16 ICRPHQI 22  
| | | | |  
Db 4 LCRPRRI 10

RESULT 11

Q25134 PRELIMINARY; PRT; 23 AA.

AC Q25134;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE HROX3 (FRAGMENT).  
GN HROX3.  
OS Haliotis rufescens (California red abalone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;  
OC Haliotidae; Haliotis.  
OX NCBI\_TaxID=6454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Degnan B.M.;  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93372986; PubMed=7689904;  
RA Degnan B.M.; Morse D.E.;  
RT "Identification of eight homeobox-containing transcripts expressed during larval development and at metamorphosis in the gastropod mollusc Haliotis rufescens."  
RT Mol. & Biol. Biotechnol. 2:1-9(1993).  
DR EMBL; X79372; CAA55917.1; -.  
DR HSSP; P02833; 9ANT.  
DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2793 MW; CC387AE7BDA6C44D CRC64;

Query Match 20.9%; Score 28; DB 5; Length 23;  
Best Local Similarity 57.1%; Pred. No. 1.3e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 16 ICRPHQI 22  
| | | | |  
Db 5 LCRPRRI 11

RESULT 12

Q78379 PRELIMINARY; PRT; 17 AA.

ID Q78379;  
AC Q78379;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE VIRAL SAMPLE FLPR5C (FLORIDA PATIENT B), PARTIAL ENV CDS, V5 REGION (FRAGMENT).  
DE Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang L.Q.; Leigh-Brown A.J.;  
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92271245; PubMed=1589796;  
RA Ou C.Y.; Ciesielski C.A.; Myers G.; Banda C.I.; Luo C.C.;  
RA Korber B.T.M.; Mullins J.I.; Schochetman G.; Berkelman R.D.;  
RA Economou A.N.; Witte J.J.; Furman L.J.; Satten G.A.; MacInnes K.A.;  
RA Curran J.W.; Jaffe H.W.;  
RT "Molecular epidemiology of HIV transmission in a dental practice."  
RL Science 256:1165-1171(1992).  
DR EMBL; M92124; AAA44494.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 17 17  
SQ SEQUENCE 17 AA; 1652 MW; 34757BBFD1240170 CRC64;

Query Match 20.1%; Score 27; DB 12; Length 17;  
Best Local Similarity 45.5%; Pred. No. 1.4e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 NNTSSTDICRP 19  
| | | | |  
Db 3 NNTNGTETFRP 13

RESULT 13

P96173 PRELIMINARY; PRT; 20 AA.

ID P96173;  
AC P96173;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE LEADER PEPTIDE.  
OS Vibrio sp. (strain 2693).  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=79682;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2693;  
RX MEDLINE=98274751; PubMed=9611817;  
RA Xu Y.; Zhang Y.; Liang Z.Y.; Van de Castele M.; Legrain C.;  
RA Glansdorff N.;  
RT "Aspartate carbamoyltransferase from a psychrophilic deep-sea

RT bacterium, Vibrio strain 2693: properties of the enzyme, genetic  
 RT organization and synthesis in *Escherichia coli*.;  
 RL Microbiology 144:1435-1441(1998).  
 DR EMBL; Y09786; CAA70922.1; -.  
 SQ SEQUENCE 20 AA; 2241 MW; 35C31F588FB5D63 CRC64;

Query Match 20.1%; Score 27; DB 2; Length 20;  
 Best Local Similarity 41.7%; Pred. No. 1.6e+03;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTTSSTDICRP 20  
 :: || : |||  
 Db 7 SLSLSPKLVPRH 18

## RESULT 14

Q9TWR5 PRELIMINARY; PRT; 20 AA.  
 AC Q9TWR5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE PNV2 TOXIN (FRAGMENT).  
 OS Phoneutria nigriventer (Brazilian armed spider).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
 OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.  
 OX NCBI\_TaxID=6918;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94030062; PubMed=8216354;  
 RA Bento A.C., Novello J.C., Marangoni S., Antunes E., Giglio J.R.,  
 RA Oliveira B., de Nucci G.;  
 RT "Identification of a new vascular smooth muscle contracting  
 RT polypeptide in *Phoneutria nigriventer* spider venom.";  
 RL Biochem. Pharmacol. 46:1092-1095(1993).  
 SQ SEQUENCE 20 AA; 2176 MW; F28C3D81D983BCA5 CRC64;

Query Match 20.1%; Score 27; DB 5; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 DICRP 19  
 |||:  
 Db 6 DICQP 10

## RESULT 15

Q91102 PRELIMINARY; PRT; 22 AA.  
 AC Q91102;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE HOX-B3-LIKE HOMEODOMAIN PROTEIN (FRAGMENT).  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID=34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BLOOD;  
 RX MEDLINE=95005122; PubMed=7921046;  
 RA Pavell A.M., Stellwag E.J.;  
 RT "Survey of Hox-like genes in the teleost *Morone saxatilis*:  
 RT implications for evolution of the Hox gene family.";  
 RL Mol. Mar. Biol. Biotechnol. 3:149-157(1994).  
 DR EMBL; U09944; AAC59650.1; -.  
 DR HSSP; P02833; 9ANT.  
 DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.  
 KW Homeobox; Nuclear protein; DNA-binding.  
 FT NON\_TER 1  
 FT NON\_TER 22  
 SQ SEQUENCE 22 AA; 2703 MW; AE4485CB7FF7CF1D CRC64;

Query Match 20.1%; Score 27; DB 13; Length 22;  
 Best Local Similarity 42.9%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 16 ICRPHQI 22  
 :||| :  
 Db 2 LCRPRRV 8

Search completed: February 12, 2002, 13:07:13  
 Job time: 177 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 13:00:40 ; Search time 23.63 Seconds  
(without alignments)  
72.098 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_185

Perfect score: 134

Sequence: 1 PCAPCTFSNTTSSDTCRPHQIC 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 213999

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_1101.\*
- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
  - 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
  - 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
  - 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
  - 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
  - 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
  - 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
  - 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
  - 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
  - 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
  - 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
  - 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
  - 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
  - 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
  - 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
  - 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
  - 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
  - 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
  - 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
  - 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
  - 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
  - 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	34.3	17	20	AAW95323
2	41	30.6	17	21	AAV51961
3	41	30.6	17	21	AAV51973
4	37	27.6	22	21	AAV51542
5	37	27.6	23	20	AAW73416
6	34.5	25.7	17	22	AAU05310
7	34.5	25.7	18	22	AAU05309
8	34.5	25.7	20	22	AAU05307
9	34.5	25.7	20	22	AAU05315
10	34.5	25.7	21	22	AAU05308
11	34.5	25.7	23	19	AAW65481

12	34	25.4	20	22	AAU05353	R1 and R2 peptide
13	33.5	25.0	20	21	AAV23019	Human APC protein
14	33	24.6	15	19	AAV20893	Human presenilin I
15	33	24.6	20	22	AAV5932	Dextranase N-termi
16	33	24.6	20	22	AAU05317	R1 and R2 peptide
17	33	24.6	20	22	AAU05346	R1 and R2 peptide
18	33	24.6	20	22	AAU05356	R1 and R2 peptide
19	33	24.6	20	22	AAU05359	R1 and R2 peptide
20	33	24.6	20	22	AAU05390	R1 and R2 peptide
21	32.5	24.3	23	6	AAV50661	Sequence of Immuno
22	32.5	24.3	23	20	AAV28912	MSF 1-alpha peptid
23	32	23.9	15	18	AAV07656	Human ATM gene pro
24	32	23.9	15	18	AAV06235	ATM epitope #1. H
25	32	23.9	15	19	AAV77164	Pharmaceutically a
26	32	23.9	15	20	AAV33126	Human umbilical co
27	32	23.9	15	22	AAV62281	ATM open reading f
28	32	23.9	16	20	AAV14391	Peptide #19 for ep
29	32	23.9	16	20	AAV15781	Antigenic peptide
30	32	23.9	17	21	AAV59437	Human delta3 fragm
31	32	23.9	20	21	AAV65710	TGF beta 2 mutant
32	32	23.9	22	18	AAV26538	Erythropoietin rec
33	31.5	23.5	18	22	AAU05323	R1 and R2 peptide
34	31.5	23.5	20	22	AAU05325	R1 and R2 peptide
35	31	23.1	10	22	AAV95723	Human complementar
36	31	23.1	15	20	AAV08939	Fibronectin-like s
37	31	23.1	15	21	AAV52487	Fibronectin-derive
38	31	23.1	18	22	AAV98822	Human cell death p
39	31	23.1	20	12	AAV15600	Immunopeptide deri
40	31	23.1	20	16	AAV84512	Hepatitis C virus
41	31	23.1	20	17	AAV91006	HCV E2 peptide E2-
42	31	23.1	20	22	AAU05360	R1 and R2 peptide
43	31	23.1	20	22	AAU05381	R1 and R2 peptide
44	30.5	22.8	20	22	AAV15022	Peptide #1456 enco
45	30.5	22.8	20	22	AAV27460	Peptide #1497 enco

ALIGNMENTS

RESULT 1

AAW95323

ID AAW95323 standard; Protein; 17 AA.

XX AC AAW95323;

XX DT 15-MAR-1999 (first entry)

XX DE Costant and variable domain sequence of C. psittaci CPS92-106.

XX KW Chlamydia; cryptic phase; elementary body phase; replicating; probenicid;

XX KW antiporphyrin acid; immune response; infection; diagnostic; assay; MOMP;

XX KW major outer membrane protein; autoimmune; inflammatory; porphyria;

XX KW Ebstein Barr virus; antioxidant.

XX OS Chlamydia psittaci.

XX PN WO9850074-A2.

XX PD 12-NOV-1998.

XX PF 06-MAY-1998; 98WO-US09237.

XX PR 18-FEB-1998; 98US-0025521.

XX PR 06-MAY-1997; 97US-0045689.

XX PR 06-MAY-1997; 97US-0045739.

XX PR 06-MAY-1997; 97US-0045779.

XX PR 06-MAY-1997; 97US-0045780.

XX PR 06-MAY-1997; 97US-0045784.

XX PR 06-MAY-1997; 97US-0045787.

XX PR 14-AUG-1997; 97US-0911593.

XX PR 18-FEB-1998; 98US-0025174.

XX PR 18-FEB-1998; 98US-0025176.

PA (UYVA-) UNIV VANDERBILT.  
 XX Mitchell WM, Stratton CW;  
 XX WPI; 1999-059653/05.  
 DR  
 XX Composition with two agents effective against different stages of  
 PT chlamydial life cycle - comprises agent targetted against cryptic  
 PT phase, against elementary body phase, against replicating phase,  
 PT probenicid and antiporphyrin  
 XX  
 PS Claim 4; Fig 3; 138pp; English.  
 XX  
 CC The invention relates to the diagnosis and management of infections by  
 CC Chlamydia species. The invention provides a composition that comprises  
 CC at least two agents, where each of the agents is effective against a  
 CC different phase of the chlamydial life cycle. The agents are selected  
 CC from: (a) agents targetted against cryptic phase of chlamydial life  
 CC cycle; (b) agents targetted against elementary body phase of chlamydial  
 CC life cycle; (c) agents targetted against replicating phase of chlamydial  
 CC life cycle; (d) probenicid, and (e) antiporphyrin acid. The composition  
 CC is used to elicit a protective immune response to Chlamydia infection in  
 CC an animal or human and is applied until the animal or human tests  
 CC negative for Chlamydia infection. It is also used to treat biological  
 CC material infected with Chlamydia. Diagnostic kits for antibody assays  
 CC against recombinant major outer membrane protein (MOMP), and for DNA  
 CC amplification assays for chlamydial genes, are used to diagnose disease,  
 CC e.g. autoimmune disease, an inflammatory disease or a disease that  
 CC occurs in an immuno-compromised individual, associated with Chlamydia  
 CC infection. The kits are used to detect chlamydial elementary bodies in a  
 CC sample. They are also used to monitor and/or modify the course of therapy  
 CC in a patient. The treatment reduces the acellular load of infectious  
 CC Epstein Barr virus. The method is also used to treat porphyria, by  
 CC reducing the number of elementary bodies and applying a drug, e.g.  
 CC cimetidine, and antioxidants, to reduce the adverse effects associated  
 CC with porphyria. Sequences AA95320 to AA95323 represent constant and  
 CC variable domain sequences of various Chlamydia species.  
 XX  
 SQ Sequence 17 AA;  
 Query Match 34.3%; Score 46; DB 20; Length 17;  
 Best Local Similarity 64.3%; Pred. No. 5;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 CAPGTFSTSTSD 15  
 Db || || || || : |  
 1 casgtasnttvaad 14  
 XX  
 RESULT 2  
 AA51961  
 ID AA51961 standard; peptide; 17 AA.  
 XX  
 AC AA51961;  
 XX  
 DT 23-JUN-2000 (first entry)  
 XX  
 DE P. pastoris lysyl oxidase peptide fragment #11.  
 XX  
 KW Lysyl oxidase; lipoxigenase; protein disulfide isomerase; phenol oxidase;  
 KW peroxidase; protein disulfide reductase; tyrosine oxidase; fodder;  
 KW sulfhydryl oxidase; food additives.  
 XX  
 OS Pichia pastoris.  
 XX  
 PN DE19840069-A1.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 03-SEP-1998; 98DE-1040069.  
 XX  
 XX 03-SEP-1998; 98DE-1040069.  
 PR

XX (BADI ) BASF AG.  
 XX Friedrich T, Bewert W, Lueddecke E, Klingler J, Heger R;  
 XX WPI; 2000-257743/23.  
 DR  
 XX Manufacture of active preparations comprises cross linking a protein,  
 PT which surrounds the active substance with an enzyme, especially a novel  
 PT lysyl oxidase from Pichia pastoris  
 XX  
 PS Claim 17; Page 17; 22pp; German.  
 XX  
 CC This invention describes a novel method to manufacture a preparation of  
 CC an active substance, where the active substance is surrounded by at  
 CC least one layer consisting of a protein that is cross-linked by an  
 CC enzyme chosen from the group of lipoxigenase, protein disulfide  
 CC isomerase, phenol oxidase and peroxidase, lysyl oxidase, protein  
 CC disulfide reductase, tyrosine oxidase or sulfhydryl oxidases. Enzymes  
 CC chosen from lipoxigenase, protein disulfide isomerase, phenol oxidase  
 CC and peroxidase, protein disulfide reductase, tyrosine oxidase or  
 CC sulfhydryl oxidases, especially lysyl oxidase are useful for formulation  
 CC of preparations of active substances. The method of the invention is  
 CC used for manufacturing preparations of active substances. The  
 CC preparations are useful as food additives or fodder or as  
 CC pharmaceuticals. AA51951-Y51962 represent fragments of the Pichia  
 CC pastoris lysyl oxidase which are used to illustrate the method of the  
 CC invention.  
 XX  
 SQ Sequence 17 AA;  
 Query Match 30.6%; Score 41; DB 21; Length 17;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 PCAPGTFSTNT 10  
 Db || || || || ||  
 7 pcapgvvynt 16  
 XX  
 RESULT 3  
 AA51973  
 ID AA51973 standard; peptide; 17 AA.  
 XX  
 AC AA51973;  
 XX  
 DT 23-JUN-2000 (first entry)  
 XX  
 DE P. pastoris lysyl oxidase fragment #11.  
 XX  
 KW Lysyl oxidase; lipoxigenase; protein disulfide isomerase; phenol oxidase;  
 KW peroxidase; protein disulfide reductase; tyrosine oxidase; food;  
 KW sulfhydryl oxidase; animal feed.  
 XX  
 OS Pichia pastoris.  
 XX  
 PN DE19840489-A1.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 04-SEP-1998; 98DE-1040489.  
 XX  
 XX 04-SEP-1998; 98DE-1040489.  
 PR  
 XX (BADI ) BASF AG.  
 XX  
 PI Friedrich T, Bewert W, Lueddecke E, Klingler J, Heger R;  
 XX  
 XX WPI; 2000-272257/24.  
 DR  
 XX Use of specified enzymes, especially lysyl oxidase, as protein  
 PT crosslinking agents for formulating compositions containing active



PT Ingredients -  
 XX Claim 17; Page 16; 20pp; German.  
 XX This invention describes a novel method where an enzyme (I) selected  
 CC from lipoygenases, protein disulfide isomerases, phenol oxidases and  
 CC peroxidases, lysyl oxidases, protein disulfide reductases, tyrosine  
 CC oxidases or sulfhydryl oxidases is used to formulate compositions  
 CC containing active ingredients. (I) is useful for crosslinking protein  
 CC layers surrounding active ingredients in food, animal feed and  
 CC pharmaceutical products. The compositions can be formulated without  
 CC using chemical crosslinking agents. AAY51963-Y51974 represent fragments  
 CC of the Pichia pastoris lysyl oxidase protein which is used to  
 CC illustrate the method of the invention.  
 XX Sequence 17 AA;  
 SQ

Query Match 30.6%; Score 41; DB 21; Length 17;  
 Best Local Similarity 70.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PCAGTFSNT 10  
 DB 7 pcagvvynt 16  
 ||||| I  
 ||||| I

RESULT 4  
 AAB51542  
 ID AAB51542 standard; Peptide; 22 AA.  
 XX  
 AC AAB51542;  
 XX  
 XX 15-FEB-2001 (first entry)  
 DT  
 DE  
 DE Yada homologous peptide #5.  
 XX  
 XX Proteobacteria; extracellular domain; virulence determinant; Yada;  
 KW adhesin; proteobacterial infection prevention; vaccine.  
 XX  
 XX Thioabacillus ferrooxidans.  
 OS  
 XX WO2000061165-A1.  
 PN  
 XX 19-OCT-2000.  
 PD  
 XX 13-APR-2000; 2000WO-US09866.  
 PF  
 XX 13-APR-1999; 99US-0129073.  
 PR  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PA  
 XX Lupas AN;  
 PI  
 XX WPI; 2000-647397/62.  
 DR  
 XX An isolated polypeptide conserved in proteobacterial extracellular  
 PT domains used in the treatment and prevention of bacterial infections -  
 PT  
 XX Example 5; Page 59; 85pp; English.  
 PS  
 XX This invention relates to peptides AAB51512 - AAB51537 which represent  
 CC conserved proteobacterial extracellular domains. Sequences  
 CC AAB51538 - AAB51618 represent peptides homologous to Yada, a yersinia  
 CC adhesin which is an important virulence determinant of the yersinia  
 CC species. The invention includes an antibody which binds to the  
 CC proteobacterial extracellular peptides, and an immunogenic composition  
 CC containing the antibody used as a vaccine to prevent infection by a  
 CC proteobacteria. The polypeptides and antibodies are useful in the  
 CC treatment and prevention of proteobacterial infections. The polypeptides  
 CC can also be used to identify compounds which antagonize the binding of a  
 CC bacterial adhesion to its ligand. The host cell can be used to produce

CC the polypeptides in a suitable culture system. The composition can be  
 CC used to vaccinate a patient against a proteobacterial infection.  
 XX  
 XX Sequence 22 AA;  
 SQ

Query Match 27.6%; Score 37; DB 21; Length 22;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;  
 QY 3 APTGFTSNTSTDICRPQI 22  
 DB 5 aptglstqt--stdavngsq 22  
 ||||| I  
 ||||| I

RESULT 5  
 AAW73416  
 ID AAW73416 standard; Protein; 23 AA.  
 XX  
 AC AAW73416;  
 XX  
 XX 19-FEB-1999 (first entry)  
 DT  
 DE Human secreted protein encoded by Gene No. 20.  
 DE  
 XX Secreted protein; human; protein therapy; gene therapy; blood disorder;  
 KW pathological condition; diagnosis; cancer; neurological disorder;  
 KW developmental abnormality; foetal deficiency; leukaemia; hepatic disease;  
 KW immune system disorder; Alzheimer's disease; cognitive disorder;  
 KW schizophrenia; prostate disease; autoimmune disorder; AIDS.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 23  
 FT /note= "unspecified amino acid"  
 FT  
 XX WO9854206-A1.  
 PN  
 XX 03-DEC-1998.  
 PD  
 XX 28-MAY-1998; 98WO-US10868.  
 PF  
 XX 29-AUG-1997; 97US-0056296.  
 PR 30-MAY-1997; 97US-0044039.  
 PR 30-MAY-1997; 97US-0048093.  
 PR 30-MAY-1997; 97US-0048101.  
 PR 30-MAY-1997; 97US-0048190.  
 PR 30-MAY-1997; 97US-0048356.  
 PR 30-MAY-1997; 97US-0050935.  
 PR 29-AUG-1997; 97US-0056250.  
 PR 29-AUG-1997; 97US-0056293.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Carter KC, Dillon PJ, Endress GA, Feng P, Ni J;  
 PI Rosen CA, Ruben SM, Yu G;  
 PI  
 XX WPI; 1999-070209/06.  
 DR N-PSDB; AAV08830.  
 DR  
 XX New isolated human genes - useful for diagnosis and treatment of,  
 PT e.g. cancers, neurological disorders, immune diseases, developmental  
 PT disorders or blood disorders  
 XX  
 XX Claim 11; Page 157; 188pp; English.  
 PS  
 XX This sequence is encoded by a cDNA of the invention, designated  
 CC Gene No. 20. This sequence represents a human secreted protein, and is  
 CC expressed ubiquitously, including T-cells and amygdala.  
 CC The DNA sequences of the invention and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions

CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the DNA sequences.  
 CC Specific uses are described for each of the DNA sequences and the encoded  
 CC proteins, based on which tissues they are most highly expressed in, and  
 CC include developing products for the diagnosis or treatment of cancer,  
 CC tumours, neurological disorders, developmental abnormalities and foetal  
 CC deficiencies, blood disorders, leukaemias, diseases of the immune system  
 CC (including allergies or asthma), hepatic disease, Alzheimer's and  
 CC cognitive disorders, schizophrenia, prostate diseases, autoimmune  
 CC disorders and AIDS. The polypeptides are also useful for identifying  
 CC their binding partners.  
 XX  
 XX Sequence 23 AA;

Query Match 27.6%; Score 37; DB 20; Length 23;  
 Best Local Similarity 43.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAPGTFSNTTSSDTC 17  
 | | | | |  
 Db 5 cgp9aagtaacsacac 20

RESULT 6  
 AAU05310  
 ID AAU05310 standard; peptide; 17 AA.  
 XX  
 AC AAU05310;  
 XX  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE R1 and R2 peptide #4 useful as erythropoietin receptor agonist.  
 XX  
 KW Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200138342-A2.  
 PN  
 XX 31-MAY-2001.  
 PD  
 XX 24-NOV-2000; 2000WO-US32224.  
 PF  
 XX 24-NOV-1999; 99US-0449064.  
 PR  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA  
 XX Balu P;  
 PI  
 XX WPI; 2001-417749/44.  
 DR  
 XX

XX The present sequence for potential R1 and R2 peptide #4 can be used  
 CC to form a peptide dimer that binds and activates the erythropoietin  
 CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
 CC (AAU05301-AAU05393) are described in the present invention. Also  
 CC described is a method for synthesizing such peptide dimers which  
 CC act as EPO-R agonists. The method is useful for synthesizing peptide  
 CC dimers which are useful, in vitro, as tools for understanding the  
 CC biological role of EPO, in the development of other compounds that bind  
 CC to EPO-R, as commercial research reagents for various medical research  
 CC and diagnostic applications, for detecting EPO receptors on living  
 CC cells, for treatment of disorders associated with a deficiency of EPO,  
 CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
 CC and chronic inflammatory diseases, autoimmune diseases and malignancies,

CC and for boosting the red blood count of a patient prior to surgery.  
 XX  
 XX Sequence 17 AA;

Query Match 25.7%; Score 34.5; DB 22; Length 17;  
 Best Local Similarity 43.8%; Pred. No. 2.1e+02;  
 Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 5 GTFS-NTTSSDTCRP 19  
 | | | | |  
 Db 1 gtyschfgpltbvorp 16

RESULT 7  
 AAU05309  
 ID AAU05309 standard; peptide; 18 AA.  
 XX  
 AC AAU05309;  
 XX  
 XX 24-OCT-2001 (first entry)  
 DT  
 DE R1 and R2 peptide #3 useful as erythropoietin receptor agonist.  
 XX  
 KW Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200138342-A2.  
 PN  
 XX 31-MAY-2001.  
 PD  
 XX 24-NOV-2000; 2000WO-US32224.  
 PF  
 XX 24-NOV-1999; 99US-0449064.  
 PR  
 XX (GLAX ) GLAXO GROUP LTD.  
 PA  
 XX Balu P;  
 PI  
 XX WPI; 2001-417749/44.  
 DR  
 XX

XX Synthesizing peptide dimer useful as erythropoietin receptor agonist by  
 CC binding linking group with functional groups as initiation sites for  
 CC peptide synthesis to support and synthesizing peptide segments -  
 XX

Example 2-5; Fig 1; 43pp; English.

XX The present sequence for potential R1 and R2 peptide #3 can be used  
 CC to form a peptide dimer that binds and activates the erythropoietin  
 CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
 CC (AAU05301-AAU05393) are described in the present invention. Also  
 CC described is a method for synthesizing such peptide dimers which  
 CC act as EPO-R agonists. The method is useful for synthesizing peptide  
 CC dimers which are useful, in vitro, as tools for understanding the  
 CC biological role of EPO, in the development of other compounds that bind  
 CC to EPO-R, as commercial research reagents for various medical research  
 CC and diagnostic applications, for detecting EPO receptors on living  
 CC cells, for treatment of disorders associated with a deficiency of EPO,  
 CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
 CC and chronic inflammatory diseases, autoimmune diseases and malignancies,  
 CC and for boosting the red blood count of a patient prior to surgery.  
 XX

XX Sequence 18 AA;

Query Match 25.7%; Score 34.5; DB 22; Length 18;  
 Best Local Similarity 43.8%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 5 GTFS-NTTSSDTCRP 19

Db 2 gtyschfgpltbvcrp 17  
 ||:| : |::|||

## RESULT 8

AAU05307  
 ID AAU05307 standard; peptide; 20 AA.

XX AC AAU05307;

XX DT 24-OCT-2001 (first entry)

XX DE R1 and R2 peptide #1 useful as erythropoietin receptor agonist.

XX KW Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.

XX OS Homo sapiens.

XX PN WO200138342-A2.

XX PD 31-MAY-2001.

XX PF 24-NOV-2000; 2000WO-US32224.

XX PR 24-NOV-1999; 99US-0449064.

XX PA (GLAX ) GLAXO GROUP LTD.

XX PI Balu P;

XX XX WPI; 2001-417749/44.

XX PT Synthesizing peptide dimer useful as erythropoietin receptor agonist by  
 PT binding linking group with functional groups as initiation sites for  
 PT peptide synthesis to support and synthesizing peptide segments -

XX PS Example 1; Fig 1; 43pp; English.

XX CC The present sequence for potential R1 and R2 peptide #1 can be used  
 CC to form a peptide dimer that binds and activates the erythropoietin  
 CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
 CC (AAU05301-AAU05333) are described in the present invention. Also  
 CC described is a method for synthesizing such peptide dimers which  
 CC act as EPO-R agonists. The method is useful for synthesizing peptide  
 CC dimers which are useful, in vitro, as tools for understanding the  
 CC biological role of EPO, in the development of other compounds that bind  
 CC to EPO-R, as commercial research reagents for various medical research  
 CC and diagnostic applications, for detecting EPO receptors on living  
 CC cells, for treatment of disorders associated with a deficiency of EPO,  
 CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
 CC and chronic inflammatory diseases, autoimmune diseases and malignancies,  
 CC and for boosting the red blood count of a patient prior to surgery.

XX SQ Sequence 20 AA;

## Query Match

Best Local Similarity 25.7%; Score 34.5; DB 22; Length 20;

Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 5 GTFS-NTTSSDTCRP 19

Db 2 gtyschfgpltbvcrp 17

## RESULT 9

AAU05315  
 ID AAU05315 standard; peptide; 20 AA.

XX AC AAU05315;

XX

DT 24-OCT-2001 (first entry)

XX DE R1 and R2 peptide #9 useful as erythropoietin receptor agonist.

XX KW Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.

XX OS Homo sapiens.

XX PN WO200138342-A2.

XX PD 31-MAY-2001.

XX PF 24-NOV-2000; 2000WO-US32224.

XX PR 24-NOV-1999; 99US-0449064.

XX PA (GLAX ) GLAXO GROUP LTD.

XX PI Balu P;

XX XX WPI; 2001-417749/44.

XX PT Synthesizing peptide dimer useful as erythropoietin receptor agonist by  
 PT binding linking group with functional groups as initiation sites for  
 PT peptide synthesis to support and synthesizing peptide segments -

XX PS Disclosure; Fig 1; 43pp; English.

XX CC The present sequence for potential R1 and R2 peptide #9 can be used  
 CC to form a peptide dimer that binds and activates the erythropoietin  
 CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
 CC (AAU05301-AAU05393) are described in the present invention. Also  
 CC described is a method for synthesizing such peptide dimers which  
 CC act as EPO-R agonists. The method is useful for synthesizing peptide  
 CC dimers which are useful, in vitro, as tools for understanding the  
 CC biological role of EPO, in the development of other compounds that bind  
 CC to EPO-R, as commercial research reagents for various medical research  
 CC and diagnostic applications, for detecting EPO receptors on living  
 CC cells, for treatment of disorders associated with a deficiency of EPO,  
 CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
 CC and chronic inflammatory diseases, autoimmune diseases and malignancies,  
 CC and for boosting the red blood count of a patient prior to surgery.

XX SQ Sequence 20 AA;

## Query Match

Best Local Similarity 25.7%; Score 34.5; DB 22; Length 20;

Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 5 GTFS-NTTSSDTCRP 19

Db 2 gtyschfgpltbvcrp 17

## RESULT 10

AAU05308

ID AAU05308 standard; peptide; 21 AA.

XX AC AAU05308;

XX DT 24-OCT-2001 (first entry)

XX DE R1 and R2 peptide #2 useful as erythropoietin receptor agonist.

XX KW Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.

XX OS Homo sapiens.

PN WO200138342-A2.  
 PD 31-MAY-2001.  
 XX  
 XX 24-NOV-2000; 2000WO-US32224.  
 XX  
 XX 24-NOV-1999; 99US-0449064.  
 XX  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Balu P;  
 XX  
 XX WPI; 2001-417749/44.  
 DR  
 XX  
 XX Synthesizing peptide dimer useful as erythropoietin receptor agonist by  
 PT binding linking group with functional groups as initiation sites for  
 PT peptide synthesis to support and synthesizing peptide segments -  
 XX  
 XX Example 10-12; Fig 1; 43pp; English.  
 PS  
 XX The present sequence for potential R1 and R2 peptide #2 can be used  
 CC to form a peptide dimer that binds and activates the erythropoietin  
 CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
 CC (AAU05301-AAU05393) are described in the present invention. Also  
 CC described is a method for synthesizing such peptide dimers which  
 CC act as EPO-R agonists. The method is useful for synthesizing peptide  
 CC dimers which are useful, in vitro, as tools for understanding the  
 CC biological role of EPO, in the development of other compounds that bind  
 CC to EPO-R, as commercial research reagents for various medical research  
 CC and diagnostic applications, for detecting EPO receptors on living  
 CC cells, for treatment of disorders associated with a deficiency of EPO,  
 CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
 CC and chronic inflammatory diseases, autoimmune diseases and malignancies,  
 CC and for boosting the red blood count of a patient prior to surgery.  
 XX  
 XX Sequence 21 AA;  
 SQ

Query Match 25.7%; Score 34.5; DB 22; Length 21;  
 Best Local Similarity 43.8%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;  
 QY 5 GTFPS-NTTSTDICRP 19  
 II:I :|||  
 Db 2 gtySCHFGPltbcrp 17.

RESULT 11  
 AAW65481  
 ID AAW65481 standard; peptide; 23 AA.  
 XX  
 AC AAW65481;  
 XX  
 XX 12-OCT-1998 (first entry)  
 DT  
 XX Hepatitis B surface antigen derived peptide (IGP 1082).  
 DE  
 XX Annexin V; hepatitis B surface antigen; immunogen; vaccine;  
 XX hepatitis delta virus; infection; HBsAg.  
 KW  
 XX Synthetic.  
 OS  
 OS Hepatitis b virus.  
 XX  
 XX WO9829442-A1.  
 PN  
 XX  
 PD 09-JUL-1998.  
 XX  
 XX 23-DEC-1997; 97WO-EP07268.  
 PF  
 XX 11-JUL-1997; 97EP-0870103.  
 PR  
 XX 30-DEC-1996; 96EP-0870164.  
 XX  
 XX (INNO-) INNOGENETICS NV.

XX De Meyer S, Depla E, Maertens G, Yap S;  
 PI WPI; 1998-388040/33.  
 XX  
 XX Immunogenic polypeptide from hepatitis B surface antigen - useful  
 PT in, e.g. vaccine against hepatitis B virus or hepatitis delta virus  
 PT infection  
 XX  
 XX Example 3; Page 35; 71pp; English.  
 PS  
 XX The invention relates to an immunogenic peptide derived from hepatitis B  
 CC surface antigen (HBsAg) which competes with the hepatitis B surface  
 CC antigen/annexin V interaction or which binds a compound or antibody  
 CC competing with the hepatitis B surface antigen/annexin V interaction.  
 CC Also claimed are: (1) a combination of the immunogenic peptide and a  
 CC negatively charged phospholipid; (2) a peptide composition comprising  
 CC the immunogenic peptide; (3) a vaccine comprising the immunogenic peptide  
 CC as an active substance; (4) antibodies which specifically bind to the  
 CC peptide and inhibit binding of HBsAg to annexin V, and (5) a therapeutic  
 CC composition comprising as an active substance the antibodies of (4).  
 CC The vaccine of (3), and the therapeutic composition of (5), can be used  
 CC as an inoculum to vaccinate humans against an infection with hepatitis  
 CC B and/or hepatitis delta virus. The immunogenic peptide can be used in a  
 CC method to detect antibodies which are capable of competing with the  
 CC hepatitis B and/or hepatitis delta virus surface antigen/annexin V  
 CC interaction. The immunogenic peptide can also be used to screen for  
 CC drugs which block the binding between annexin V and the peptide, and as  
 CC a therapeutic to treat humans infected with hepatitis B virus and/or  
 CC hepatitis delta virus. The present sequence represents one of the  
 CC peptide fragments derived from HBsAg which were synthesised to map the  
 CC annexin V-binding site on HBsAg.  
 XX  
 XX Sequence 23 AA;  
 SQ

Query Match 25.7%; Score 34.5; DB 19; Length 23;  
 Best Local Similarity 44.4%; Pred. No. 2.8e+02;  
 Matches 8; Conservative 3; Mismatches 4; Indels 3; Gaps 1;  
 QY 1 PCAPGTFSTNTSTDICR 18  
 I III :||| I;  
 Db 1 pllpgt---sttstgpc 15

RESULT 12  
 AAU05353  
 ID AAU05353 standard; peptide; 20 AA.  
 XX  
 AC AAU05353;  
 XX  
 XX 24-OCT-2001 (first entry)  
 DT  
 XX R1 and R2 peptide #47 useful as erythropoietin receptor agonist.  
 DE  
 XX Human; erythropoietin receptor; EPO-R; EPO agonist; EPO deficiency;  
 KW renal failure; anaemia; chronic inflammatory disease;  
 KW autoimmune disease; malignancy; red blood count.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200138342-A2.  
 PN  
 XX 31-MAY-2001.  
 PD  
 XX 24-NOV-2000; 2000WO-US32224.  
 PF  
 XX 24-NOV-1999; 99US-0449064.  
 PR  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Balu P;  
 PI  
 XX

DR WPI; 3001-417749/44.  
XX  
PT Synthesizing peptide dimer useful as erythropoietin receptor agonist by  
PT binding linking group with functional groups as initiation sites for  
PT peptide synthesis to support and synthesizing peptide segments .  
XX  
XX  
PS Disclosure; Fig 1; 43pp; English.  
XX  
XX  
CC The present sequence for potential R1 and R2 peptide #47 can be used  
CC to form a peptide dimer that binds and activates the erythropoietin  
CC receptor (EPO-R). Various possible peptide sequences for R1 and R2  
CC (AAU05301-AAU05393) are described in the present invention. Also  
CC described is a method for synthesizing such peptide dimers which  
CC act as EPO-R agonists. The method is useful for synthesizing peptide  
CC dimers which are useful, in vitro, as tools for understanding the  
CC biological role of EPO, in the development of other compounds that bind  
CC to EPO-R, as commercial research reagents for various medical research  
CC and diagnostic applications, for detecting EPO receptors on living  
CC cells, for treatment of disorders associated with a deficiency of EPO,  
CC such as end-stage renal failure/dialysis, anaemia associated with AIDS  
CC and chronic inflammatory diseases, autoimmune diseases and malignancies,  
CC and for boosting the red blood count of a patient prior to surgery.  
XX  
XX Sequence 20 AA:  
  
Query Match 25.4%; Score 34; DB 22; Length 20;  
Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 14 TDICRPH 20  
Db 12 tbverpn 18  
  
RESULT 13  
AAB23019  
ID AAB23019 standard; peptide; 20 AA.  
XX  
XX AAB23019;  
AC  
XX  
XX  
DT 16-JAN-2001 (first entry)  
XX  
DE Human APC protein 20 aa repeat #4 (1643-1662).  
XX  
XX APC gene; Adenomatous Polyposis Coli gene; human; chromosome 5q21;  
KW familial adenomatous polyposis; FAP locus; Gardner's syndrome; GS;  
KW sporadic tumour; adenoma; carcinoma; cancer; lung; breast; colon; rectum;  
KW bladder; liver; sarcoma; stomach; prostate; cancer; lymphoma;  
KW tumour suppressor; anti-APC antibody; detection; diagnosis; prognosis;  
KW genetic predisposition; drug screening; DP2.5; repeat region.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US6114124-A.  
PN  
XX  
XX 05-SEP-2000.  
PD  
XX  
XX 25-MAY-1995; 95US-0450582.  
XX  
XX 16-JAN-1991; 91GB-0000962.  
PR  
XX 16-JAN-1991; 91GB-0000963.  
PR  
XX 16-JAN-1991; 91GB-0000974.  
PR  
XX 16-JAN-1991; 91GB-0000975.  
PR  
XX 08-AUG-1991; 91US-0741940.  
PR  
XX 12-AUG-1994; 94US-0289548.  
PR  
XX (ICIL ) IMPERIAL CHEM IND PLC.  
PA (UYJO ) UNIV JOHNS HOPKINS.  
PA (UTAH ) UNIV UTAH.  
PA (CANC-) CANCER INST.  
XX  
XX Carlson M, Groden J, Joslyn G, Kinzler K, Markham AF, Anand R;  
PI

PI Albertsen H, White RL, Thliveris A, Nakamura Y, Vogelstein B;  
XX Hedge PJ;  
XX  
DR WPI; 2000-565003/52.  
XX  
XX  
PT Detecting Adenomatous Polyposis Coli (APC) protein in a sample for  
PT diagnosing cancers, involves contacting the sample with antibodies that  
PT specifically bind to APC protein and detecting the complex formed .  
XX  
XX  
PS Example 15; Column 33-34; 125pp; English.  
XX  
XX The invention relates to a novel method for detecting Adenomatous  
XX Polyposis Coli (APC) protein in a sample. The method involves  
XX contacting the sample with antibodies which specifically binds to the  
XX 2843 amino acid form of the human APC protein, or to a mutant APC  
XX protein, and detecting an APC-antibody complex. Mutations in the APC  
XX gene play a role in tumorigenesis, indicating that it is a tumour  
XX suppressor gene. It is located on chromosome 5q21, which corresponds to  
XX the FAP (familial adenomatous polyposis) locus. FAP is an autosomal  
XX dominant inherited disease in which affected individuals develop  
XX hundreds to thousands of adenomatous polyps in the colon and rectum,  
XX some of which progress to malignancy. The FAP locus is often found to  
XX be deleted in sporadic (i.e., non-familial) adenomas and carcinomas, and  
XX chromosome 5q deletions have also been observed in tumours of the lung,  
XX breast, colon, rectum, bladder, liver, sarcomas, stomach, and prostate,  
XX and in leukaemias and lymphomas. Although the FAP locus contains  
XX several other genes such as FER, TBI, TB2, and MCC, it is thought that  
XX mutations in the APC gene play a key role in the development of FAP and  
XX sporadic tumours. The method is useful for detecting APC protein and its  
XX mutant forms in foetal tissue, placental tissue, amniotic fluid, blood,  
XX serum or a tumour sample. The method is useful for diagnosing or  
XX prognosing neoplastic tissue, for detecting a genetic predisposition to  
XX cancer, for detecting germline and somatic alteration of wild-type APC  
XX genes, and for testing therapeutic agents for the ability to suppress  
XX tumours. Sequences AAB23016-B23022 represent seven 20 amino acid repeats  
XX that are semiregularly spaced in the human APC protein.  
XX  
XX Sequence 20 AA:  
  
Query Match 25.0%; Score 33.5; DB 21; Length 20;  
Best Local Similarity 44.4%; Pred. No. 3.4e+02;  
Matches 8; Conservative 2; Mismatches 5; Indels 3; Gaps 1;  
  
QY 2 CAPGT---FSNTTSSTDI 16  
Db 2 cvegtptinfstatslsdl 19  
  
RESULT 14  
AAY20893  
ID AAY20893 standard; Protein; 15 AA.  
XX  
XX AAY20893;  
AC  
XX  
XX 22-JUL-1999 (first entry)  
DT  
XX  
XX Human presenilin I mutant protein fragment 39.  
DE  
XX  
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
KW frameshift mutation; age-related disease; neurodegenerative disorder;  
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-W;  
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
KW bcl-2; B-cell leukaemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
KW high mobility group protein-C; neuroendocrine specific protein A.  
XX  
XX Synthetic.  
OS  
XX Homo sapiens.  
XX

PN WO9845322-A2.  
XX  
PD 15-OCT-1998.  
XX  
PF 02-APR-1998; 98WO-IB00705.  
XX  
XX 10-APR-1997; 97US-0043163.  
PR  
XX (UYUT-) RIJKSNIV UTRECHT.  
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
XX  
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
XX  
XX WPI; 1998-609901/51.  
DR N-PSDB; AAX75761.  
DR  
XX  
PT Diagnosing disease by detecting frameshift mutations in RNA or  
PT corresponding protein mutations - used to diagnose cancer and  
PT neurological diseases, particularly Alzheimer's disease, and also  
PT for treatment and prevention with specific ribozymes or wild-type  
PT RNA  
XX  
PS Disclosure; Figure 10; 258pp; English.  
XX  
CC This invention describes a novel method for the diagnosis of a disease  
CC caused by, or associated with, an RNA molecule that has a frameshift  
CC mutation. The method is used to diagnose age-related diseases, especially  
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
CC and many others listed) or susceptibility to these disorders. The method  
CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
CC at an early stage. It is based on the observation that disease may be  
CC caused by mutations in RNA rather than DNA. The invention describes the  
CC use of neuronal system RNA molecules, specifically proteins including  
CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
CC proteins Tau and Big Tau, ubiquitin-B, apolipoprotein E, microtubule  
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic  
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
CC 2 (bcl-2) proto-oncogene, semaphorin III, HSPF-1, high mobility group  
CC protein-C (HMGP-C) and neuroendocrine specific protein A.  
XX  
SQ Sequence 15 AA;  
  
Query Match 24.6%; Score 33; DB 19; Length 15;  
Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 CARGTF 7  
Db | | | |  
4 ccpgtf 9  
  
RESULT 15  
AAR75932  
ID AAR75932 standard; Protein; 16 AA.  
XX  
AC AAR75932;  
XX  
DT 24-FEB-1996 (first entry)  
XX  
DE Dextranase N-terminal sequence.  
XX  
KW Dextranase; thermostable enzyme; dextran hydrolysis; sugar.  
XX  
OS Penicillium minioluteum HI-4.  
XX  
PN EP663443-A1.  
XX  
PD 19-JUL-1995.

XX 13-DEC-1994; 94EP-0203614.  
PF  
XX 14-DEC-1993; 93CU-0000115.  
PR  
XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.  
PA  
XX Delgado Boada JM, Campana HR, Curbelo DM, Jimenez ER;  
PI Patron CF, Cremata Alvarez JA, Garcia Fernandez R;  
PI Garcia Garcia BM, Gonzalez Martinez ME, Herrera Martinez LS;  
PI Margollez Clark E, Morera Cordova V, Raices Perez-Castaneda MR;  
XX  
DR WPI; 1995-247530/33.  
XX  
XX Nucleotide sequence encoding Penicillin minioluteum dextranase -  
PT useful for the high-level industrial prodn. of dextranase.  
PT  
XX  
PS Example 2; Page 7; 29pp; English.  
XX  
CC The N-terminal region of P. minioluteum dextranase was sequenced  
CC using the Edman degradation method. Dextranase may be secreted as  
CC a recombinant protein by the transformed host, Pichia pastoris.  
CC The recombinant enzyme has higher thermal stability than the natural  
CC P. minioluteum enzyme (specifically an optimum temp. of 55-60 deg  
CC and a half-life of 7.6 hr at 50 deg), and it can be used in the sugar  
CC industry for sugarcane juice dextran hydrolysis.  
XX  
SQ Sequence 16 AA;  
  
Query Match 24.6%; Score 33; DB 16; Length 16;  
Best Local Similarity 46.2%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
  
QY 5 GTFSNTTSSSTDIC 17  
Db | | : | | |  
2 gttntnhtcgadfc 14  
  
Search completed: February 12, 2002, 13:03:52  
Job time: 192 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 13:02:40 ; Search time 12.45 seconds  
(without alignments)  
41.572 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_185

Perfect score: 134

Sequence: 1 PCAGPTFSNTTSTSDICRPHQIC 23

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 119836

Minimum DB seq length: 0

Maximum DB seq length: 23

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	53.7	15	1 US-08-221-583-45	Sequence 45, Appl
2	72	53.7	15	1 US-08-221-583-46	Sequence 46, Appl
3	72	53.7	15	5 PCT-US95-04018-45	Sequence 45, Appl
4	72	53.7	15	5 PCT-US95-04018-46	Sequence 46, Appl
5	71	53.0	15	1 US-08-221-583-44	Sequence 44, Appl
6	71	53.0	15	5 PCT-US95-04018-44	Sequence 44, Appl
7	57	42.5	15	1 US-08-221-583-47	Sequence 47, Appl
8	57	42.5	15	5 PCT-US95-04018-47	Sequence 47, Appl
9	48	35.8	20	2 US-08-126-016-24	Sequence 24, Appl
10	44	32.8	15	1 US-08-221-583-48	Sequence 48, Appl
11	44	32.8	15	5 PCT-US95-04018-48	Sequence 48, Appl
12	39	29.1	23	3 US-09-101-146-58	Sequence 58, Appl
13	33	24.6	16	1 US-08-354-618-3	Sequence 3, Appl
14	33	24.6	20	2 US-08-934-915-141	Sequence 141, App
15	32	23.9	15	1 US-08-493-092-4	Sequence 4, Appl
16	32	23.9	15	1 US-08-508-836A-4	Sequence 4, Appl
17	32	23.9	15	2 US-08-629-001A-4	Sequence 4, Appl
18	32	23.9	15	4 US-08-642-274D-4	Sequence 4, Appl
19	32	23.9	15	4 US-08-952-127-4	Sequence 4, Appl
20	32	23.9	15	4 US-08-952-014C-4	Sequence 4, Appl
21	32	23.9	22	1 US-08-484-635-178	Sequence 178, App
22	32	23.9	22	2 US-08-484-631-178	Sequence 178, App
23	32	23.9	22	2 US-08-827-570-178	Sequence 178, App
24	31.5	23.5	20	1 US-08-318-193-63	Sequence 63, Appl
25	31	23.1	18	4 US-09-461-697-449	Sequence 449, App
26	31	23.1	18	6 5217891-2	Patent No. 5217891
27	31	23.1	20	2 US-08-934-915-78	Sequence 78, Appl

28	31	23.1	20	4 US-08-612-973-83	Sequence 83, Appl
29	31	23.1	20	4 US-08-927-597-83	Sequence 83, Appl
30	30	22.4	18	1 US-08-323-531-8	Sequence 8, Appl
31	30	22.4	18	1 US-08-198-094-8	Sequence 8, Appl
32	30	22.4	18	4 US-08-107-794A-8	Sequence 8, Appl
33	30	22.4	18	5 PCT-US93-07424-8	Sequence 8, Appl
34	30	22.4	18	5 PCT-US95-02087-8	Sequence 8, Appl
35	30	22.4	19	1 US-08-484-635-244	Sequence 244, App
36	30	22.4	19	2 US-08-484-631-244	Sequence 244, App
37	30	22.4	19	2 US-08-827-570-244	Sequence 244, App
38	29.5	22.0	16	1 US-08-574-763-7	Sequence 7, Appl
39	29	21.6	6	4 US-08-750-142B-20	Sequence 20, Appl
40	29	21.6	9	1 US-08-054-860-14	Sequence 14, Appl
41	29	21.6	9	3 US-08-442-378-14	Sequence 14, Appl
42	29	21.6	12	4 US-08-602-999A-280	Sequence 280, App
43	29	21.6	15	1 US-08-221-583-49	Sequence 49, Appl
44	29	21.6	15	4 US-08-602-999A-315	Sequence 315, App
45	29	21.6	15	5 PCT-US95-04018-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1  
US-08-221-583-45  
; Sequence 45, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-45

Query Match 53.7%; Score 72; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00083;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSNTTSTSDICRP 19  
|||||  
Db 1 GTFSNTTSTSDIARP 15

RESULT 2



US-08-221-583-46  
; Sequence 46, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-46

Query Match 53.7%; Score 72; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00083;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 SNTTSSTDICRPHQI 22  
||||||| |||||  
DB 1 SNTTSSTDICRPHQI 15

RESULT 3  
PCT-US95-04018-45  
; Sequence 45, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018

; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-45

Query Match 53.7%; Score 72; DB 5; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00083;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSNTTSSTDICRP 19  
||||||| |||||  
DB 1 GTFSNTTSSTDICRP 15

RESULT 4  
PCT-US95-04018-46  
; Sequence 46, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:

NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04018-46

Query Match 53.7%; Score 72; DB 5; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.00083;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 SNTTSSDTCRPHQI 22  
Db 1 SNTTSSDTCRPHQI 15

RESULT 5  
US-08-221-583-44  
; Sequence 44, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-44

Query Match 53.0%; Score 71; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APGTFSTNTSSTDI 16  
Db 2 APGTFSTNTSSTDI 15

RESULT 6  
PCT-US95-04018-44  
; Sequence 44, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-44

Query Match 53.0%; Score 71; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APGTFSTNTSSTDI 16  
Db 2 APGTFSTNTSSTDI 15

RESULT 7  
US-08-221-583-47  
; Sequence 47, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595rls  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia

STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,583  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0185  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-221-583-47

Query Match 42.5%; Score 57; DB 1; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.082;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQI 22  
Db 1 TSSTDICRPHQI 12

RESULT 8  
PCT-US95-04018-47  
Sequence 47, Application PC/TUS9504018  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
APPLICANT: Kruszynski, Marian  
APPLICANT: Mervic, Miljenko  
APPLICANT: Weber, Robert W.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04018  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,580  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,583  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,581  
FILING DATE: 01-APR-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Deluca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04018-47

Query Match 42.5%; Score 57; DB 5; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.082;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQI 22  
Db 1 TSSTDICRPHQI 12

RESULT 9  
US-08-126-016-24  
Sequence 24, Application US/08126016  
Patent No. 5811261  
GENERAL INFORMATION:  
APPLICANT: WALLACH, DAVID  
APPLICANT: NOPAR, YARON  
APPLICANT: KEMPER, OLIVER  
APPLICANT: ENGELMANN, HARTMUT  
APPLICANT: BRAKEBUSCH, CORD  
APPLICANT: ADERKA, DAN  
TITLE OF INVENTION: EXPRESSION OF THE RECOMBINANT TUMOR  
NECROSIS FACTOR BINDING PROTEIN I (TBP-I)  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Neimark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/126,016  
FILING DATE: 24-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/625668  
FILING DATE: 13-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, ROGER L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: WALLACH4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-126-016-24

Query Match 35.8%; Score 48; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCAPGTF 8  
| | | | | | | |  
Db 12 PCAPGTF 19

RESULT 10

US-08-221-583-48  
; Sequence 48, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221.583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-221-583-48

Query Match 32.8%; Score 44; DB 1; Length 15;  
Best Local Similarity 88.9%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 TDICRPHQI 22  
| | | | | | | |  
Db 1 TDIARPHQI 9

RESULT 11

PCT-US95-04018-48  
; Sequence 48, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US95-04018-48

Query Match 32.8%; Score 44; DB 5; Length 15;  
Best Local Similarity 88.9%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 TDICRPHQI 22  
| | | | | | | |  
Db 1 TDIARPHQI 9

RESULT 12

US-09-101-146-58  
; Sequence 58, Application US/09101146  
; Patent No. 6124125  
; GENERAL INFORMATION:  
; APPLICANT: Dartmouth College, St. Vincents Institute of  
; APPLICANT: Medical Research, Kemp et al.  
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: 66 E. Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: WINDOWS 95  
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/101,146  
; FILING DATE: October 7, 1998

CLASSIFICATION: 435  
PRIOR APPLICATION DATA: PN7450  
APPLICATION NUMBER: 8 JAN 1996  
FILING DATE: 8 JAN 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: DC-0050  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (856) 810-1515  
TELEFAX: (856) 810-1454  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-09-101-146-58

Query Match 29.1%; Score 39; DB 3; Length 23;  
Best Local Similarity 43.8%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 CAPGTFSTNTSSDTC 17  
||| : | : |  
Db 4 CAAGTTTGATGTATC 19

RESULT 13  
US-08-354-618-3  
; Sequence 3, Application US/08354618  
; Patent No. 5637491  
; GENERAL INFORMATION:  
; APPLICANT: Campana, Hernan Roca  
; APPLICANT: Garcia, Blanca Maria Garcia  
; APPLICANT: Clark, Emilio Margollez  
; APPLICANT: Curbelo, Dania Mateu  
; APPLICANT: Boada, Julio Marcos Delgado  
; APPLICANT: Martinez, Luis S. Herrera  
; APPLICANT: Alvarez, Jos Alberto Cremata  
; APPLICANT: Perez-Casta eda, Manuel Rafael Raices  
; APPLICANT: Martinez, Maria Elena Gonz lez  
; APPLICANT: Jim nez, Efrain Rodriguez  
; TITLE OF INVENTION: Dextranase enzyme, method for its  
; TITLE OF INVENTION: production and DNA encoding the enzyme  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ronald J. Baron  
; ADDRESSEE: Hoffmann & Baron  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: diskette - 3.5 inch, 1.44 Mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 6.0 version B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/354,618  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CU 115/93  
; FILING DATE: 14-December-1993  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO

US-08-354-618-3

Query Match 24.6%; Score 33; DB 1; Length 16;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSTNTSSDTC 17  
|| : || : |  
Db 2 GTTNTHCGADFC 14

RESULT 14  
US-08-934-915-141  
; Sequence 141, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWEE-MING  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. Foutch  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-934-915-141

Query Match 24.6%; Score 33; DB 2; Length 20;  
Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 SNTTSSTDCIRPH 20  
|| : || : |  
Db 7 SNEVSSPEIRQH 19

RESULT 15  
US-08-493-092-4  
; Sequence 4, Application US/08493092

; Patent No., 5728807  
; GENERAL INFORMATION:  
; APPLICANT: Shiloh, Yosef  
; APPLICANT: Taglie, Danilo A.  
; APPLICANT: Collins, Francis S.  
; TITLE OF INVENTION: Ataxia-Telangiectasia Gene  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Reising, Ethington, Barnard & Perry  
; STREET: P.O. Box 4390  
; CITY: Troy  
; STATE: Michigan  
; COUNTRY: US  
; ZIP: 48099  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/493,092  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kohn, Kenneth I.  
; REGISTRATION NUMBER: 30,955  
; REFERENCE/DOCKET NUMBER: P-310 (TAV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (810) 689-3500  
; TELEFAX: (810) 689-4071  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-493-092.4

Query Match 23.9%; Score 32; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 8 SNTTSSTDIC 17  
I: : |||:|  
Db 6 SSASQSTDLC 15

Search completed: February 12, 2002, 13:04:30  
Job time: 110 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: February 12, 2002, 12:54:49 ; Search time 13.27 Seconds  
(Without alignments)  
223.874 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_201  
Perfect score: 216  
Sequence: 1 PCAPGTFSTSTSDICRPHQICNVVAIPGNASMDAVCT 39  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 8589  
Minimum DB seq length: 0  
Maximum DB seq length: 39  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR,f8:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38.5	17.8	37	2 G49050	T-cell surface gly
2	38	17.6	33	2 A37479	huwentoxin-I - Chi
3	34	15.7	37	2 S68261	hypothetical prote
4	33.5	15.5	39	2 A05323	phospholipase A2 (
5	33	15.3	39	2 G82613	hypothetical prote
6	32.5	15.0	35	2 A60959	agelenin - funnel-
7	32	14.8	30	2 A22977	delta-endotoxin -
8	31	14.4	20	2 A42865	Ca2+/calmodulin-de
9	31	14.4	30	2 S70343	napin large chain
10	31	14.4	34	2 E49410	t-complex polypept
11	31	14.4	39	2 A60716	somatotropin intro
12	30.5	14.1	38	2 A48158	pheromone precursor
13	30	13.9	18	2 B49048	T-cell receptor be
14	30	13.9	26	2 S28994	antifungal protein
15	30	13.9	29	2 C61233	conceptus protein
16	30	13.9	35	2 A39830	dentin matrix, inc
17	30	13.9	38	2 A45495	beta-defensin-1 -
18	29	13.4	13	2 S47381	T-cell antigen rec
19	29	13.4	16	2 S38292	30K allergen - rye
20	29	13.4	17	2 A60317	glucagon-like pept
21	29	13.4	21	2 I54351	gene HEXA protein
22	29	13.4	22	2 A39269	LX-1 tumor antigen
23	29	13.4	24	2 S08293	lectin - sunn hemp
24	29	13.4	26	2 A42865	Ca2+/calmodulin-de
25	29	13.4	31	2 S53233	gene X protein - h
26	29	13.4	32	2 I48415	heat shock factor
27	29	13.4	34	2 D81044	hypothetical prote
28	28	13.0	18	1 A58589	alpha-conotoxin EI
29	28	13.0	23	2 S60565	homeodomain protei

30	28	13.0	25	2 S74094	fibulin 1 variant
31	28	13.0	27	2 S28995	antifungal protein
32	28	13.0	27	2 C44636	homeotic protein H
33	28	13.0	27	2 T12330	metallothionein -
34	28	13.0	29	1 GCFUE	glucagon - Europea
35	28	13.0	29	2 A61135	glucagon - bigeye
36	28	13.0	30	1 TIPUIW	trypsin inhibitor
37	28	13.0	30	2 A47607	immunogenic protei
38	28	13.0	31	2 S21743	trypsin inhibitor
39	28	13.0	32	2 E82089	hypothetical protei
40	27.5	12.7	30	2 S28991	antifungal protein
41	27.5	12.7	31	2 A55430	conotoxin NgVIA -
42	27.5	12.7	35	2 PL0164	alpha-lactalbumin
43	27.5	12.7	36	2 A82208	hypothetical prote
44	27.5	12.7	36	2 JT0513	Ig heavy chain V-I
45	27	12.5	20	2 A39328	notechis II-5b non

ALIGNMENTS

RESULT 1

G49050  
T-cell surface glycoprotein CD8 beta-1 chain, secreted form 5 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-1994 #sequence\_revision 03-Nov-1995 #text\_change 29-Aug-1997  
C:Accession: G49050  
R:DiSanto, J.P.; Smith, D.; de Bruin, D.; Lacy, E.; Flomenberg, N.  
Eur. J. Immunol. 23, 320-326, 1993  
A:Title: Transcriptional diversity at the duplicated human CD8 beta loci.  
A:Reference number: A49050; MUID:93170376  
A:Accession: G49050  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-37 <DIS>  
A:Note: sequence extracted from NCBI backbone (NCBIP:125543)  
C:Genetics:  
A:Gene: GDB:CD8B1; CD8B  
A:Cross-references: GDB:119771; OMIM:186730  
A:Map position: 2p12-2p12  
C:Keywords: alternative splicing; extracellular protein; glycoprotein

RESULT 2

A37479  
huwentoxin-I - Chinese bird spider  
C:Species: Selenocosmia huwena (Chinese bird spider)  
C:Date: 18-Mar-1994 #sequence\_revision 07-Oct-1994 #text\_change 16-Feb-1996  
C:Accession: A37479; JCI089  
R:Liang, S.P.; Zhang, D.Y.; Pan, X.; Chen, Q.; Zhou, P.A.  
Toxicol 31, 969-978, 1993  
A:Title: Properties and amino acid sequence of huwentoxin-I, a neurotoxin purified fr  
A:Reference number: A37479; MUID:94024948  
A:Accession: A37479  
A:Molecule type: protein  
A:Residues: 1-33 <LIA>  
R:Liang, S.P.; Zong, X.; Luo, J.C.; Jing, H.; Gu, X.C.  
Acta Sci. Natur. Univ. Pekin. 29, 668-674, 1993  
A:Title: Secondary structure study of huwentoxin-I, a neurotoxin from the venom of th  
A:Reference number: JCI089  
A:Accession: JCI089  
A:Molecule type: protein  
A:Residues: 1-33 <LI2>  
C:Comment: This peptide is the major active protein component of venom in this specie

C:Keywords: presynaptic neurotoxin: venom  
F:2-17,9-22,16-29/Disulfide bonds: #status experimental

Query Match 17.6%; Score 38; DB 2; Length 33;  
Best Local Similarity 30.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 5 GTFSNTSSTDICRPHQICN 24  
| | | : | | | |  
Db 4 GVFDACPGKNECCPNRVC5 23

RESULT 3  
S68261  
hypothetical protein gadd7.2 - long-tailed hamster  
C:Species: Cricetus longicaudatus (long-tailed hamster)  
C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S68261  
R:Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.  
Nucleic Acids Res. 24, 1589-1593, 1996  
A:Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth, but lacks  
A:Reference number: S68260; MUID:96211359  
A:Accession: S68261  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-37 <HOL>  
A:Cross-references: EMBL:L40430

Query Match 15.7%; Score 34; DB 2; Length 37;  
Best Local Similarity 60.0%; Pred. No. 7.9e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CNVVAIPGNA 32  
| | | | | | | |  
Db 10 CTVTRAPGNA 19

RESULT 4  
A05323  
phospholipase A2 (EC 3.1.1.4) - Mexican beaded lizard (fragment)  
N:Alternate names: phosphatidylcholine 2-acylhydrolase  
C:Species: Heloderma horridum (Mexican beaded lizard)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 05-Apr-1995  
C:Accession: A05323

R:Sosa, B.P.; Alagon, A.C.; Martin, B.M.; Possani, L.D.  
Biochemistry 25, 2927-2933, 1986  
A:Reference number: A05323; MUID:86243292  
A:Note: H. h. horridum  
A:Accession: A05323  
A:Molecule type: protein  
A:Residues: 1-39 <SOS>  
C:Comment: Phospholipases A2 catalyze the calcium-dependent hydrolysis of the 2-acyl glycerol  
C:Superfamily: phospholipase A2  
C:Keywords: calcium; carboxylic ester hydrolase; lipid degradation; venom

Query Match 15.5%; Score 33.5; DB 2; Length 39;  
Best Local Similarity 31.0%; Pred. No. 9.6e+02;  
Matches 9; Conservative 3; Mismatches 10; Indels 7; Gaps 2;

QY 2 CAPGTFSN-----TTSSTDI-CRPHQIC 23  
| | | : | | | | | |  
Db 11 CGAGNAADYSQLGTEKTDMDCCRDHDC 39

RESULT 5  
G82613  
hypothetical protein xfl988 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82613

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: G82613  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-39 <SIM>  
A:Cross-references: GB:AE004018; GB:AE003849; NID:g9107093; PIDN:AAF84790.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreras-Neto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frachado, M.A.; Madeira, A.M.B.N.; Madeira, E.L.; Kitajima, J.P.; Kieger, J.E.; Kuramae, E.E.; La Chado, M.A.; Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, R.; Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XFI988

Query Match 15.3%; Score 33; DB 2; Length 39;  
Best Local Similarity 37.5%; Pred. No. 1.1e+03;  
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 TFSNTTSTDICRPHQ 21  
: | | | | | | | |  
Db 24 SLSKVTVSADVMRAHR 39

RESULT 6  
A60959  
agelenin - funnel-weaving spider (Agelena opulenta)  
C:Species: Agelena opulenta  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 24-Feb-1994  
C:Accession: A60959  
R:Hagiwara, K.; Sakai, T.; Miwa, A.; Kawai, N.; Nakajima, T.  
Biomed. Res. 11, 181-186, 1990

A:Title: Complete amino acid sequence of a new type of neurotoxin from the venom of  
A:Reference number: A60959  
A:Accession: A60959  
A:Molecule type: protein  
A:Residues: 1-35 <HAG>  
C:Keywords: calcium channel inhibitor; presynaptic neurotoxin; venom

Query Match 15.0%; Score 32.5; DB 2; Length 35;  
Best Local Similarity 42.9%; Pred. No. 1.2e+03;  
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 17 CRPH-QTCNVVAIP 29  
| | | : | | | |  
Db 3 CLPHNRCNALSGP 16

RESULT 7  
A22977  
delta-endotoxin - Bacillus thuringiensis (fragment)  
C:Species: Bacillus thuringiensis  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 07-Feb-1997  
C:Accession: A22977  
R:Armstrong, J.L.; Rohrmann, G.F.; Beaudreau, G.S.  
J. Bacteriol. 161, 39-46, 1985  
A:Reference number: A22977; MUID:85104736  
A:Accession: A22977



```

RESULT 10
E9410
t-complex polypeptide 1 homolog (peak 2 fraction) - rabbit (fragments)
N:Alternate names: chaperonin homolog (peak 2)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 04-Sep-1998
C:Accession: E9410
R:Rommlaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.;
  proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
A:Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven r
A:Reference number: A49410; MUID:94089752
A:Accession: E9410
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-34 <ROW>
A:Experimental source: reticulocyte
A:Note: sequence modified after extraction from NCBI backbone
A:Note: sequence extracted from NCBI backbone (NCBIP:141043)
C:Superfamily: molecular chaperone t-complex-type

Query Match 14.4%; Score 31; DB 2; Length 34;
Best Local Similarity 38.9%; Pred. No. 1.8e+03;
Matches 7; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 19 PHQICNVVAIPGNASMDA 36
      I I I:
      I I I:
DB 17 PRQLCD-----NAGFDA 28

RESULT 11
A60716
somatotropin intron-related protein RDE.25 - rat (fragment)
N:Alternate names: growth hormone gene-related protein RDE.25
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999
C:Accession: A60716
R:Montpetit, M.L.; Tenniswood, M.P.
  J. Cell. Biochem. 39, 285-292, 1989
A:Title: Does the lack of regression-associated mRNA expression render a rat ventral
A:Reference number: A60716; MUID:89214371
A:Accession: A60716
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-39 <MON>
C:Comment: The mRNA encoding this hypothetical protein shows homology at the nucleoti
he complete mRNA may encode a protein related to somatotropin through abnormal splici

Query Match 14.4%; Score 31; DB 2; Length 39;
Best Local Similarity 27.6%; Pred. No. 2e+03;
Matches 8; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 2 CAPCFTFSNTTSSTDICRPHQICNVVAIPG 30
      I I I: I: I: I: I:
DB 10 CLPEVLSSIPSTHVAHHLQFLNLPSSG 38

RESULT 12
A48158
pheromone precursor MF alpha - fungus (Filobasidium floriforme)
C:Species: Filobasidiella neoformans, Cryptococcus neoformans
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: A48158
R:Moore, T.D.; Edman, J.C.
  Mol. Cell. Biol. 13, 1962-1970, 1993
A:Title: The alpha-mating type locus of Cryptococcus neoformans contains a peptide ph
A:Reference number: A48158; MUID:93180845
A:Accession: A48158
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-38 <MOO>
A:Note: sequence inconsistent with nucleotide translation

```

A; Note: sequence extracted from NCBI backbone (NCBIN:126064, NCBIPI:126070)

Query Match 14.1%; Score 30.5; DB 2; Length 38;  
Best Local Similarity 35.7%; Pred. No. 2.2e+03;  
Matches 10; Conservative 4; Mismatches 9; Indels 5; Gaps 1;

QY 7 FSNTSTSDICRPHQICNVVAIPGNASM 34  
|||::| |::|  
Db 12 FSAATSSSEAPR-----NQEAHPGGMTL 34

## RESULT 13

B49048  
T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997  
C:Accession: B49048  
R:Sloud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.  
Eur. J. Immunol. 22, 2413-2418, 1992  
A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juvenile  
A:Reference number: A49048; MUID:92387250  
A:Accession: B49048  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-18 <SIO>  
A:Experimental source: patient EV, IL-2R+ synovial T-cells  
A:Note: sequence extracted from NCBI backbone (NCBIPI:113264)  
C:Keywords: T-cell receptor

Query Match 13.9%; Score 30; DB 2; Length 18;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 CAPGTFSTNTT 11  
||||:|  
Db 7 CAPGXYGYT 16

## RESULT 14

S28994  
antifungal protein 2 - white mustard (fragment)  
C:Species: Sinapis alba (white mustard)  
C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 09-Jun-2000  
C:Accession: S28994  
R:Terras, F.R.G.; Torrekens, S.; van Leuven, F.; Osborn, R.W.; Vanderleyden, J.; Cammue, F.E.B. Lett. 316, 233-240, 1993  
A:Title: A new family of basic cysteine-rich plant antifungal proteins from Brassicaceae  
A:Reference number: S28989; MUID:93138130  
A:Accession: S28994  
A:Molecule type: protein  
A:Residues: 1-26 <TER>  
C:Superfamily: gamma-thionin  
C:Keywords: phosphoprotein

Query Match 13.9%; Score 30; DB 2; Length 26;  
Best Local Similarity 35.7%; Pred. No. 1.8e+03;  
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 GTFSTSTSDICR 18  
||:|::|  
Db 9 GTWSGVCGNNACR 22

## RESULT 15

C61233  
conceptus protein 5 - cat (fragment)  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: C61233  
R:Thatcher, M.J.D.; Shille, V.M.; Fliss, M.F.; Bazer, F.W.; Sisum, W.; Randal, S.

Biol. Reprod. 44, 108-120, 1991  
A:Title: Characterization of feline conceptus proteins during pregnancy.  
A:Reference number: A61233; MUID:91198359

A:Accession: C61233  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-29 <THA>  
C:Superfamily: insulin-like growth factor binding protein 1; thyroglobulin type I rep

Query Match 13.9%; Score 30; DB 2; Length 29;  
Best Local Similarity 20.0%; Pred. No. 2e+03;  
Matches 7; Conservative 5; Mismatches 11; Indels 12; Gaps 1;

QY 2 CAPGTFSTSTSDICRPHQICNVVAIPGNASMDA 36  
||||:|::|  
Db 7 CAP-----CSPEKALCPVPDSCQTQSA 29

Search completed: February 12, 2002, 12:56:43  
Job time: 114 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:56:25 ; Search time 10.18 Seconds  
(without alignments)  
140.464 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_201

Perfect score: 216

Sequence: 1 PCAPGTFSTNTSSDTCRPHQICNVAILPGNSMDAVCT 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 2657

Minimum DB seq length: 0

Maximum DB seq length: 39

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	17.6	33	1 TXH1_SELHU	P56676 selenocosmi
2	37.5	17.4	37	1 TXJC_HADVE	P82228 hadronyche
3	37	17.1	36	1 TXJA_HADVE	P82227 hadronyche
4	34	15.7	36	1 TXJB_HADVE	P82226 hadronyche
5	33.5	15.5	39	1 PA2_HELHO	P04362 heloderma h
6	32.5	15.0	35	1 TXAG_AGEOP	P31328 agelena opu
7	31	14.4	28	1 PPOX_BOVIN	P56602 bos taurus
8	31	14.4	30	1 CY01_VIOOD	P82230 viola odora
9	30	13.9	31	1 CYUA_PSYLO	P56872 psychotria
10	30	13.9	38	1 BD01_BOVIN	P46159 bos taurus
11	29	13.4	15	1 DIDH_PSESP	P80701 pseudomonas
12	29	13.4	24	1 LEC_CROJU	P16352 crotalaria
13	29	13.4	34	1 RR2_OCHNE	Q40606 ochrosphaer
14	28	13.0	18	1 CXAL_CONER	P50382 conus ermin
15	28	13.0	29	1 GLUC_PLAHE	P23062 platichthys
16	28	13.0	30	1 ITI1_LAGLE	P26771 lagenaria l
17	28	13.0	30	1 ITI1_CITLA	P11969 citrullus l
18	28	13.0	36	1 GLU1_ORENI	P81026 oreochromis
19	27.5	12.7	31	1 CXD6_CONNI	P56710 conus nigro
20	27	12.5	22	1 CXM1_CONGE	P01523 conus geogr
21	27	12.5	23	1 AFP2_BRANA	P03026 brassica na
22	27	12.5	24	1 RS13_THETH	P80377 thermus aqu
23	27	12.5	25	1 ANDT_ANDAU	P56684 androctonus
24	27	12.5	27	1 AFP1_BRARA	P30227 brassica ra
25	27	12.5	28	1 ETX2_BACCE	P80568 bacillus ce
26	27	12.5	28	1 PAXC_PSEPO	P20260 pseudechis
27	27	12.5	30	1 VPU_HVISC	P05948 human immun
28	27	12.5	30	1 LPTV_ECOLI	P03061 escherichia
29	26.5	12.3	32	1 IAPP_SHEEP	Q28605 ovis aries
30	26	12.0	38	1 NLFX_WHEAT	P39085 triticum ae
31	26	12.0	30	1 GLUM_ANGAN	P41521 anguilla an
32	26	12.0	30	1 ITRI_MOMCH	P10294 momordica c
33	26	12.0	33	1 VTI1_RAT	P58200 rattus norv

34 26 12.0 34 1 DEF2\_RABIT P07468 oryctolagus  
35 26 12.0 37 1 RL36\_MYCPN P52864 mycoplasma  
36 25.5 11.8 23 1 CP23\_SPOER P56683 spodoptera  
37 25.5 11.8 23 1 PAP2\_SPOEX P30256 spodoptera  
38 25.5 11.8 23 1 PAP3\_SPOEX P30257 spodoptera  
39 25.5 11.8 32 1 CAPP\_METEX Q49136 methylobact  
40 25.5 11.8 34 1 TX1\_SCOGR P56855 scodra grls  
41 25.5 11.8 37 1 RL36\_LEPIN O9kd13 leptospira  
42 25.5 11.8 37 1 RL36\_VIRCH P78001 vibrio chol  
43 25.5 11.8 37 1 TXM2\_AGEAP P11058 agelenopsis  
44 25.5 11.8 38 1 BD08\_BOVIN P46166 bos taurus  
45 25 11.6 11 1 TKNA\_ONCMY P28499 oncorhynchu

## ALIGNMENTS

RESULT 1

TXH1\_SELHU  
ID TXH1\_SELHU STANDARD; PRT; 33 AA.  
AC P56676; Q9NJC2;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HUWENTOXIN-I (HWTX-I)  
OS Selenocosmia huwena (Chinese bird spider).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
OC Mygalomorphae; Theraphosidae; Selenocosmia.  
OX NCBI\_TaxID=29017;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=94024948; Pubmed=8212049;  
RA Liang S.-P., Zhang D.-Y., Pan X., Chen Q., Zhou P.-A.;  
RT "Properties and amino acid sequence of huwentoxin-I, a neurotoxin  
RT purified from the venom of the Chinese bird spider Selenocosmia  
RT huwena.";  
RL Toxicon 31:969-978(1993).  
RN [2]  
RP SEQUENCE OF 1-32 FROM N.A.  
RC STRAIN=Huwena;  
RA Li M., Zhou Z., Liang S.;  
RT "Huwentoxin-I (HWTX-I) peptide cDNA sequence.";  
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP STRUCTURE BY NMR.  
RC TISSUE=Venom;  
RX MEDLINE=97408601; Pubmed=9263120;  
RA Qu Y.-X., Liang S.-P., Ding J., Liu X.-C., Zhang R.-J., Gu X.-C.;  
RT "Proton nuclear magnetic resonance studies on huwentoxin-I from the  
RT venom of the spider Selenocosmia huwena: 2. Three-dimensional  
RT structure in solution.";  
RL J. Protein Chem. 16:565-574(1997).  
RN [4]  
RP DISULFIDE BONDS.  
RC TISSUE=Venom;  
RX MEDLINE=94183409; Pubmed=8136023;  
RA Zhang D.-Y., Liang S.-P.;  
RT "Assignment of the three disulfide bridges of huwentoxin-I, a  
RT neurotoxin from the spider Selenocosmia huwena.";  
RL J. Protein Chem. 12:735-740(1993).  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE=97179771; Pubmed=9028007;  
RA Zhou P.-A., Xie X.-J., Li M., Yang D.-M., Xie Z.-P., Zong X.,  
RA Liang S.-P.;  
RT "Blockade of neuromuscular transmission by huwentoxin-I, purified from  
RT the venom of the Chinese bird spider Selenocosmia huwena.";  
RL Toxicon 35:39-45(1997).  
CC -I- FUNCTION: LETHAL NEUROTOXIN. BINDS TO THE NICOTINIC ACETYLCHOLINE  
CC RECEPTOR. BLOCKS NEUROMUSCULAR TRANSMISSION.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



```

RT amide form, and investigation of the disulfide bond arrangement.";
RL Biomed. Res. 12:357-363(1991).
[3]
RN
RP SYNTHESIS, DISULFIDE BONDS, AND AMIDATION.
RC TISSUE-Venom;
RX MEDLINE-93043890; PubMed-1421801;
RA Inui T., Hagiwara K., Nakajima K., Kimura T., Nakajima T.,
RA Sakakibara S.;
RT "Synthesis and disulfide structure determination of agelatin:
RT identification of the carboxy-terminus as an amide form.";
RC Pept. Res. 5:140-144(1992).
CC -!- FUNCTION: SUPPRESSES THE EXCITATORY POSTSYNAPTIC POTENTIALS
CC -!- POSSIBLY BY BLOCKING THE PRESYNAPTIC CALCIUM CHANNEL.
CC FIR; A60959; A60959.
DR Venom; Neurotoxin; Calcium channel inhibitor; Amidation.
KW DISULFID 3 19
FT DISULFID 10 24
FT DISULFID 18 34
FT DISULFID 35 35
FT MOD_RES 35 AA; 3825 MW; CBE6462825350D90 CRC64;
FT SEQUENCE 35 AA; 3825 MW; CBE6462825350D90 CRC64;
Query Match 15.0%; Score 32.5; DB 1; Length 35;
Best Local Similarity 42.9%; Pred. No. 4e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 1; Gaps
QY 17 CRPH-QICNVWVAP 29
DB 3 CLPHNFCNALSGP 16
| | | | | | | |
| | | | | | | |
QY 17 CRPH-QICNVWVAP 29
DB 3 CLPHNFCNALSGP 16
| | | | | | | |
| | | | | | | |
RESULT 7
PPOX BOVIN
ID PPOX_BOVIN STANDARD; PRT; 28 AA.
AC P56602;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO) (FRAGMENTS).
GN PPOX
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE-95331315; PubMed-7607249;
RA Taketani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,
RA Nishimura K., Inokuchi H.;
RT "Induction of terminal enzymes for heme biosynthesis during
RT differentiation of mouse erythroleukemia cells.";
RL Eur. J. Biochem. 230:760-765(1995).
CC -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-
CC IX + H2O(2).
CC -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW Flavoprotein; FAD; Mitochondrion.
FT NON_TER 1 1
FT NON_CONS 13 14
FT NON_TER 28 28
FT SEQUENCE 28 AA; 2777 MW; FEDFC3F09CB6A345 CRC64;
Query Match 14.4%; Score 31; DB 1; Length 28;

```

```

Best Local Similarity 42.9%; Pred. No. 5e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 6; Conservative 3;

Qy 25 VVAIPGNASMDAVC 38
    ||:|:|:|:|
Db 4 VVVLGGIGSDSLC 17

RESULT 8
CYOL_VIOOD
ID CYOL_VIOOD STANDARD; PRT; 30 AA.
AC P82230;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CYCLOVIOOLACIN O1.
OS Viola odorata (Sweet violet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Violaceae; Viola.
OX NCBI_TaxID=97441;
RN [1]
RP SEQUENCE.
RX MEDLINE=20069951; PubMed=10600388;
RA Craik J.D., Daly N.D., Bond T., Wayne C.;
RT "Plant cyclotides - a unique family of cyclic and knotted proteins
RT that defines the cyclic cysteine knot structural motif.";
RL J. Mol. Biol. 294:1327-1336(1999).
CC -|- PTM: THIS IS A CYCLIC PEPTIDE.
CC -|- SIMILARITY: BELONGS TO THE CYCLOTIDE FAMILY.
CC -|- CAUTION: THIS PEPTIDE BEING CYCLIC, ITS SEQUENCE WAS ARBITRARILY
CC CHOSEN TO START AT THE POSITION SHOWN BELOW. AS THE DNA SEQUENCE
CC FROM WHICH IT IS DERIVED IS NOT YET KNOWN, IT IS NOT POSSIBLE TO
CC ASSIGN THE CORRECT N- AND C-TERMINI.
FT DISULFID 2 17
FT DISULFID 7 22
FT DISULFID 15 28
SQ SEQUENCE 30 AA; 3141 MW; A3D8CA231098E7E7 CRC64;

Query Match 14.4%; Score 31; DB 1; Length 30;
Best Local Similarity 37.5%; Pred. No. 5.4e+02;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 23 CNVVAIPGNASMDAVC 38
    ||:|:|:|:|
Db 7 CTVTALLGSCSNVC 22

RESULT 9
CYLA_PSYLO
ID CYLA_PSYLO STANDARD; PRT; 31 AA.
AC P56872; P82254;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CYCLOPSYCHORIDE A (CPT).
OS Psychotria longipes.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Rubiaceae; Psychotria.
OX NCBI_TaxID=41680;
RN [1]
RP SEQUENCE.
RX MEDLINE=95230294; PubMed=7714530;
RA Witherup K.M., Bogusky M.J., Anderson P.S., Ramjit H., Ransom R.W.,
RA Wood T., Sardana M.;
RT "Cyclopsychotride A, a biologically active, 31-residue cyclic peptide
RT isolated from Psychotria longipes.";
RL J. Nat. Prod. 57:1619-1625(1994).
RN [2]
RP SYNTHESIS, AND ANTIBACTERIAL ACTIVITY.
RX MEDLINE=99362685; PubMed=10430870;

```

```

Tam J.P., Lu Y.-A., Yang J.-L., Chiu K.-W.;
RT "An unusual structural motif of antimicrobial peptides containing
RT end-to-end macrocycle and cystine-knot disulfides.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:8913-8918(1999).
CC -|- FUNCTION: PROBABLY PARTICIPATES IN A PLANT DEFENSE MECHANISM. HAS
CC ANTIBIOTIC ACTIVITY. INHIBITS NEUTROPHIL BINDING. ACTIVE AGAINST
CC BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
CC -|- PTM: THIS IS A CYCLIC PEPTIDE.
CC -|- SIMILARITY: BELONGS TO THE CYCLOTIDE FAMILY.
CC -|- CAUTION: THIS PEPTIDE BEING CYCLIC, ITS SEQUENCE WAS ARBITRARILY
CC CHOSEN TO START AT THE POSITION SHOWN BELOW. AS THE DNA SEQUENCE
CC FROM WHICH IT IS DERIVED IS NOT YET KNOWN, IT IS NOT POSSIBLE TO
CC ASSIGN THE CORRECT N- AND C-TERMINI.
KW Antibiotic.
FT DISULFID 2 17 BY SIMILARITY.
FT DISULFID 7 22 BY SIMILARITY.
FT DISULFID 15 29 BY SIMILARITY.
SQ SEQUENCE 31 AA; 3255 MW; ACC0BBB232ED0CD0 CRC64;

Query Match 13.9%; Score 30; DB 1; Length 31;
Best Local Similarity 37.5%; Pred. No. 7.5e+02;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 23 CNVVAIPGNASMDAVC 38
    ||:|:|:|:|
Db 7 CTVTALLGSCSKVC 22

RESULT 10
BD01_BOVIN
ID BD01_BOVIN STANDARD; PRT; 38 AA.
AC P46159;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE BETA-DEFENSIN 1 (BNDB-1) (BNBD-1).
GN DEFB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX STRAIN=HEREFORD; TISSUE=Neutrophils;
RX MEDLINE=93203264; PubMed=8454635;
RA Selsted M.E., Tang Y.-Q., Morris W.L., McGuire P.A., Novotny M.J.,
RA Smith W., Henschen A.H., Cullor J.S.;
RT "Purification, primary structures, and antibacterial activities of
RT beta-defensins, a new family of antimicrobial peptides from bovine
RT neutrophils.";
RL J. Biol. Chem. 268:6641-6648(1993).
CC -|- FUNCTION: HAS BACTERICIDAL ACTIVITY. ACTIVE AGAINST E. COLI ML35
CC BUT NOT AGAINST S. AUREUS 502A.
CC -|- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC -|- TISSUE SPECIFICITY: NEUTROPHILIC GRANULES.
CC HSSP; P46170; 1BNB.
DR InterPro: IPR001271; Defensin.
DR InterPro: IPR001855; Defensin_beta.
DR Pfam: PF00711; Defensin_beta; 1.
DR SMART: SM00048; DEFSN; 1.
KW Antibiotic.
FT DISULFID 5 34 BY SIMILARITY.
FT DISULFID 12 27 BY SIMILARITY.
FT DISULFID 17 35 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4278 MW; 48B872D1025E1A68 CRC64;

Query Match 13.9%; Score 30; DB 1; Length 38;
Best Local Similarity 21.9%; Pred. No. 9.2e+02;
Matches 7; Conservative 8; Mismatches 11; Indels 6; Gaps 1;

```

QY 7 FNTSTSDICRPHQICNVVAIPGNASMDAVC 38  
 IDH\_PESP  
 DB 2 FASCHNGGICLPNR-----CPGHMIQIGIC 27

RESULT 11  
 DIDH\_PESP  
 ID DIDH\_PESP STANDARD; PRT; 15 AA.  
 AC P80701;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE 3-ALPHA-HYDROXYSTEROID DEHYDROGENASE (EC 1.1.1.50) (3-ALPHA-HSD)  
 DE (HYDROXYPROSTAGLANDIN DEHYDROGENASE) (HSD29) (FRAGMENT).  
 OS Pseudomonas sp.  
 CC Bacteria; Proteobacteria.  
 OX NCBI\_TaxID=306;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=97100200; PubMed=8944761;  
 RA Oppermann U.C.T., Maser E.;  
 RT "Characterization of a 3 alpha-hydroxysteroid dehydrogenase/carbonyl  
 reductase from the gram-negative bacterium Comamonas testosteroni.";  
 RL Eur. J. Biochem. 241:744-749(1996).  
 CC -!- FUNCTION: ALONG WITH THE 3 ALPHA-HYDROXYSTEROID DEHYDROGENASE AND  
 A/B RING STEROIDS, IT ALSO REDUCES SEVERAL XENOBIOTIC CARBONYL  
 COMPOUNDS, INCLUDING A METRAPONE-BASED CLASS OF INSECTICIDES, TO  
 THE RESPECTIVE ALCOHOL METABOLITES.  
 CC -!- CATALYTIC ACTIVITY: ANDROSTERONE + NAD(P)(+) =  
 5-ALPHA-ANDROSTANE-3,17-DIONE + NAD(P)H.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 (SDR) FAMILY.  
 CC InterPro: IPR002198; ADH\_short.  
 DR PROSITE: PS00061; ADH\_SHORT; PARTIAL.  
 KW Oxidoreductase; NAD.  
 FT DOMAIN 6 >15 INVOLVED IN COFACTOR BINDING  
 FT NON\_TER 15 15  
 FT SEQUENCE 15 AA; 1315 MW; 9506860D070A7790 CRC64;

Query Match 13.4%; Score 29; DB 1; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 4.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 25 VVAIPGNAS 33  
 IDH\_PESP  
 DB 2 VIAITGSAS 10

RESULT 12  
 LEC\_CROJU  
 ID LEC\_CROJU STANDARD; PRT; 24 AA.  
 AC P16352;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE LECTIN (FRAGMENT).  
 OS Crotalaria juncea (Sunn hemp).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Crotalariaeae;  
 CC Crotalaria.  
 OX NCBI\_TaxID=3829;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RA Foriers A., de Neve R., Strosberg A.D.;  
 RT "Lectin sequences as a tool for chemotaxonomical classification.";  
 RL Physiol. Veg. 17:597-606(1979).  
 CC -!- SUBUNIT: HOMOTETRAMER.

CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR PIR; S08293; S08293.  
 DR HSP; P04122; ILGB.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; PARTIAL.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; PARTIAL.  
 KW Lectin; Glycoprotein.  
 FT NON\_TER 24 24  
 FT SEQUENCE 24 AA; 2614 MW; 80704D8CD9F9BB6E CRC64;

Query Match 13.4%; Score 29; DB 1; Length 24;  
 Best Local Similarity 60.0%; Pred. No. 7.9e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TFSNTSTSD 15  
 IDH\_PESP  
 DB 5 SFSSTKFTSD 14

RESULT 13  
 RR2\_OCHNE  
 ID RR2\_OCHNE STANDARD; PRT; 34 AA.  
 AC Q40606;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S2 (FRAGMENT).  
 GN RPS2.  
 OS Ochrosphaera neapolitana.  
 CC Chloroplast.  
 CC Eukaryota; Haptophyceae; Isochrysidales; Ochrosphaera.  
 OX NCBI\_TaxID=35137;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCMP 593;  
 RA Huss V.A.R., Tietze A.C., Julius C.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE S2p FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; X99078; CAA67534.1; -  
 DR InterPro: IPR001865; Ribosomal\_S2.  
 DR Pfam: PF00318; Ribosomal\_S2; 1.  
 DR PROSITE: PS00962; RIBOSOMAL\_S2\_1; PARTIAL.  
 DR PROSITE: PS00963; RIBOSOMAL\_S2\_2; PARTIAL.  
 KW Ribosomal protein; Chloroplast.  
 FT NON\_TER 1 1  
 FT SEQUENCE 34 AA; 3640 MW; 2039BA0FB5710655 CRC64;

Query Match 13.4%; Score 29; DB 1; Length 34;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 24 NVVAIPGNASMDAV 37  
 IDH\_PESP  
 DB 1 NLVDIPIPANDDAI 14

RESULT 14  
 CXAL\_CONER  
 ID CXAL\_CONER STANDARD; PRT; 18 AA.  
 AC P50982;



```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ALPHA-CONOTOXIN EI.
OS Conus ermineus (Atlantic fish-hunting cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=55423;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=96062516; PubMed=7578057;
RA Martinez J.S., Olivera B.M., Gray W.R., Craig A.G., Groebe D.R.,
RA Abramson S.N., McIntosh J.M.;
RT "Alpha-conotoxin EI, a new nicotinic acetylcholine receptor
RT antagonist with novel selectivity.";
RL Biochemistry 34:14519-14526(1995).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; Hydroxylation.
FT DISULFID 4 10
FT DISULFID 5 18
FT MOD_RES 3 3 HYDROXYLATION.
FT MOD_RES 18 18 AMIDATION.
SQ SEQUENCE 18 AA; 2082 MW; 60A61A6C427A6B5E CRC64;

Query Match 13.0%; Score 28; DB 1; Length 18;
Best Local Similarity 45.5%; Pred. No. 8e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 15 DICRPHQICNV 25
Db | | | | |
2 DPCCYHPTCNM 12

RESULT 15
GLUC_PLAFE
ID GLUC_PLAFE STANDARD; PRT; 29 AA.
AC P23062;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE GLUCAGON.
OS Platichthys flesus (European flounder), and
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Platichthys.
OX NCBI_TaxID=8260, 8241;
RN [1]
RP SEQUENCE.
RC SPECIES=P.flesus;
RX MEDLINE=87219793; PubMed=3556313;
RA Conlon J.M., Davis M.S., Thim L.;
RT "Primary structure of insulin and glucagon from the flounder
RT (Platichthys flesus).";
RL Gen. Comp. Endocrinol. 66:203-209(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=T.obesus; TISSUE=Pancreas;
RX MEDLINE=92009094; PubMed=1916209;
RA Navarro I., Gutierrez J., Caixach J., Rivera J., Planas J.;
RT "Isolation and primary structure of glucagon from the endocrine
RT pancreas of Thunnus obesus.";
RL Gen. Comp. Endocrinol. 83:227-232(1991).
CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC THE BLOOD SUGAR LEVEL.
CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.

```

```

CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; S09348; GCFLF.
DR PIR; A61135; A61135.
DR HSSP; P01274; IGCN.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; Hormone2; 1.
DR PRINTS; PR00275; GLUCAGON.
DR SMART; SM00070; GLUCA; 1.
DR PROSITE; PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone.
SQ SEQUENCE 29 AA; 3508 MW; 77D5943208662E52 CRC64;

Query Match 13.0%; Score 28; DB 1; Length 29;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GTFSNNTS 12
Db | | | | |
4 GTFSNDYS 11

```

Search completed: February 12, 2002, 12:59:28  
Job time: 183 sec

---

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	121	56.0	30	4	Q9UIH1	Q9uih1 homo sapien	
2	41	19.0	36	5	Q9BP43	Q9bp43 conus penna	
3	38.5	17.8	37	4	Q9UDB4	Q9udb4 homo sapien	
4	36	16.7	35	12	Q70293	Q70293 human immun	
5	36	16.7	35	12	Q70294	Q70294 human immun	
6	36	16.7	35	12	Q70295	Q70295 human immun	
7	36	16.7	35	12	Q70351	Q70351 human immun	
8	35	16.2	26	9	Q9ZXH9	Q9zxh9 bacteriopho	
9	34.5	16.0	36	6	P79330	P79330 bos taurus	
10	34	15.7	33	4	Q9BY50	Q9bys0 homo sapien	
11	34	15.7	33	12	Q9IEX1	Q9ilex1 cotton leaf	
12	33	15.3	15	11	Q9QUY5	Q9quy5 rattus sp.	
13	33	15.3	39	2	Q9PBZ7	Q9pbz7 xyella fas	
14	32	14.8	33	4	Q9UD12	Q9ud12 homo sapien	
15	32	14.8	34	12	Q9QL92	Q9ql92 human adeno	
16	32	14.8	39	12	Q99IP1	Q99ip1 human immun	
17	32	14.8	39	12	Q99IP0	Q99ip0 human immun	
18	32	14.8	29	12	Q99IN9	Q99in9 human immun	
19	31.5	14.6	29	6	Q9MZ88	Q9mz88 pongo pygma	

```

OS Conus pennaceus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
RT "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL; AF215130; AAG60535.1; -.
FT NON_TER 1
SQ SEQUENCE 36 AA; 3683 MW; 04C74E40FF71141C CRC64;

Query Match 19.0%; Score 41; DB 5; Length 36;
Best Local Similarity 30.8%; Pred. No. 85;
Matches 8; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 3 APTGFTSNTSSTDICRP----HQICN 24
| | | | | : | | | | | : | |
Db 7 ASGKFMVNLKQPLCCPFGGCHLCH 32
| | | | | : | | | | | : | |

RESULT 3
Q9UDB4 PRELIMINARY; PRT; 37 AA.
ID Q9UDB4
AC Q9UDB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CD8 BETA CHAIN ISOFORM S BETA5 (FRAGMENT).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93170376; PubMed=8436166;
RA Disanto J.P., Smith D., de Bruin D., Lacy E., Flomenberg N.;
RT "Transcriptional diversity at the duplicated human CD8 beta loci.";
RL Eur. J. Immunol. 23:320-326(1993).
SQ SEQUENCE 37 AA; 3940 MW; 7A4AE1C901E15E99 CRC64;

Query Match 17.8%; Score 38.5; DB 4; Length 37;
Best Local Similarity 34.8%; Pred. No. 2e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 1 PCAPGTFSTSTSDICRPHQIC 23
| | | | | : | | | | | : | |
Db 13 PLSPNACMDTDA---ILQPHRSC 32
| | | | | : | | | | | : | |

RESULT 4
Q70293 PRELIMINARY; PRT; 35 AA.
ID Q70293
AC Q70293;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).

```

```

RT of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93329178; PubMed=8335967;
RA Cheingsong-Popov R., Bobkov A.F., Garaev M., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
their distribution in the Commonwealth of Independent States (Former
Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10705; AAA19262.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 12; Length 35;
Best Local Similarity 47.6%; Pred. No. 4.2e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTSTST--DICRPH 20
| | | | | : | | | | | : | |
Db 14 CAPGQAFYTTGEIGDIRRAH 34
| | | | | : | | | | | : | |

RESULT 5
Q70294 PRELIMINARY; PRT; 35 AA.
ID Q70294
AC Q70294;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93329178; PubMed=8335967;
RA Cheingsong-Popov R., Bobkov A.F., Garaev M., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
their distribution in the Commonwealth of Independent States (Former
Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10706; AAA19263.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 12; Length 35;
Best Local Similarity 47.6%; Pred. No. 4.2e+02;

```

```

Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTNTSST--DICRPH 20
    ||||  ||  ||||
Db 14 CAPGQAFYTTGTEIGDIRRAH 34

RESULT 6
ID Q70295 PRELIMINARY; PRT; 35 AA.
AC Q70295;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;
RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
RT of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93329178; PubMed=8335967;
RA Cheingsong-Popov R., Bobkov A.F., Garaev M., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
RT their distribution in the Commonwealth of Independent States (Former
RT Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
RT analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10707; AAA19264.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 12; Length 35;
Best Local Similarity 47.6%; Pred. No. 4.2e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTNTSST--DICRPH 20
    ||||  ||  ||||
Db 14 CAPGQAFYTTGTEIGDIRRAH 34

RESULT 7
ID Q70351 PRELIMINARY; PRT; 35 AA.
AC Q70351;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN, V3 REGION (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94338597; PubMed=8060542;
RA Bobkov A.F., Garaev M., Rzhabinova A., Kaleebu P., Pitman R.,
RA Weber J.N., Cheingsong-Popov R.;

```

```

RT "Molecular epidemiology of HIV-1 in the former Soviet Union: analysis
RT of env V3 sequences and their correlation with epidemiologic data.";
RL AIDS 8:619-624(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93329178; PubMed=8335967;
RA Cheingsong-Popov R., Bobkov A.F., Garaev M., Kaleebu P., Callow D.,
RA Rzhabinova A., Saukhat S.R., Burdajev N.P., Kolomijets N.D.,
RA Weber J.N.;
RT "Identification of human immunodeficiency virus type 1 subtypes and
RT their distribution in the Commonwealth of Independent States (Former
RT Soviet Union) by serologic V3 peptide-binding assays and V3 sequence
RT analysis.";
RL J. Infect. Dis. 168:292-297(1993).
DR EMBL; U10764; AAA19329.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3852 MW; 5C5827FCBD5DB873 CRC64;

Query Match 16.7%; Score 36; DB 12; Length 35;
Best Local Similarity 47.6%; Pred. No. 4.2e+02;
Matches 10; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

QY 2 CAPGTFSTNTSST--DICRPH 20
    ||||  ||  ||||
Db 14 CAPGQAFYTTGTEIGDIRRAH 34

RESULT 8
Q9ZXKH9 PRELIMINARY; PRT; 26 AA.
ID Q9ZXKH9 PRELIMINARY; PRT; 26 AA.
AC Q9ZXKH9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE PUTATIVE VERTEX HEAD SUBUNIT (FRAGMENT).
GN G24.
OS Bacteriophage ARI.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=66711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ARI.
RA Syu W.J., Ding H.C., Seah J.N., Wu K.M., Yu S.L., Tam M.F.,
RA Chang Y.C.;
RT "Characterization of ARI coliphage specific to enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022930; AAD01756.1; -.
FT NON_TER 26
FT NON_TER 26
SQ SEQUENCE 26 AA; 2830 MW; 5875E0CFBB665934 CRC64;

Query Match 16.2%; Score 35; DB 9; Length 26;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 8 SNTTSSSTDICRPHQICNVVA 27
    ||||  ||  ||||
Db 11 STTNSNSIGRP----NLVA 26

RESULT 9
P79330 PRELIMINARY; PRT; 36 AA.
ID P79330 PRELIMINARY; PRT; 36 AA.
AC P79330;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

```

```

DT 01-NOV-1998 (TrEMBLrel.. 08, Last annotation update)
DE AGGRCAN EPIDERMAL GROWTH FACTOR-LIKE DOMAIN-1 (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CARTILAGE PROTEOGLYCAN;
RX MEDLINE=97079270; PubMed=8921002;
RA Fulop C., Cs-Szabo G., Glant T.T.;
RT "Species-specific alternative splicing of the epidermal growth factor-
RT like domain I of cartilage aggrecan.";
RL Biochem. J. 319:935-940(1996).
DR EMBL; L29486; AAB48067.1; -.
FT NON_TER 1
SQ SEQUENCE 36 AA; 3614 MW; 9F6C32F15D3BC9DB CRC64;

Query Match 16.0%; Score 34.5; DB 6; Length 36;
Best Local Similarity 34.8%; Pred. No. 6.9e+02;
Matches 8; Conservative 1; Mismatches 7; Indels 7; Gaps 1;

QY 1 PCAPGTFSTNTSSTDCRPHQIC 23
||: || | | |
Db 9 PCSAGTCQETEG-----HWIC 24

RESULT 10
Q9BYSO PRELIMINARY; PRT; 33 AA.
AC Q9BYSO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE KERATIN ASSOCIATED PROTEIN 2.1B (FRAGMENT).
GN KRTAP2.1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCALP;
RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.;
RT "Characterization of a cluster of human high/ ultrahigh keratin
RT associated proteins on chromosome 17q12-21.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ406929; CAC27568.1; -.
FT NON_TER 1
SQ SEQUENCE 33 AA; 3420 MW; 1DAC91CBB7120302 CRC64;

Query Match 15.7%; Score 34; DB 4; Length 33;
Best Local Similarity 41.2%; Pred. No. 7.5e+02;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNTSSTDC 17
|| | | | | |
Db 17 PCGQTPCSTTCRTSSC 33

RESULT 11
Q9IEX1 PRELIMINARY; PRT; 33 AA.
AC Q9IEX1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE AC4 PROTEIN (FRAGMENT).
GN AC4.
OS cotton leaf curl virus.

```

```

OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=53010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P12-IR;
RX MEDLINE=20318672; PubMed=10859391;
RA Sanz A.I., Fraile A., Garcia-Arenal F., Zhou X., Robinson D.J.,
RA Khalid S., Butt T., Harrison B.D.;
RT "Multiple infection, recombination and genome relationships among
RT begomovirus isolates found in cotton and other plants in Pakistan.";
RL J. Gen. Virol. 81:1839-1849(2000).
DR EMBL; AJ270854; CAB97069.1; -.
DR InterPro; IPR002488; Geminiv_C4.
DR Pfam; PF01492; Geminiv_C4; 1.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3478 MW; F3121B92E34ED31E CRC64;

Query Match 15.7%; Score 34; DB 12; Length 33;
Best Local Similarity 41.7%; Pred. No. 7.5e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 NTTSTSDICRPH 20
|: | | | |
Db 15 NSNAGTTVLRPH 26

RESULT 12
Q9QUY5 PRELIMINARY; PRT; 15 AA.
AC Q9QUY5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE: CERAMIDE GALACTOSYLTRANSFERASE
DE (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=96085162; PubMed=8521863;
RA Schulte S., Stoffel W.;
RT "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate
RT transporter. Copurification, separation and characterization of the
RT two glycoproteins.";
RL Eur. J. Biochem. 233:947-953(1995).
SQ SEQUENCE 15 AA; 1657 MW; 84474749A06BFFCC CRC64;

Query Match 15.3%; Score 33; DB 11; Length 15;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PGTFSTNTS 12
|| | | | |
Db 2 PGTFSTNTS 10

RESULT 13
Q9PBZ7 PRELIMINARY; PRT; 39 AA.
AC Q9PBZ7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN XF1988.
GN XF1988.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=9A5C;
RX  MEDLINE=20365717; PubMed=10910347;
RA  Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA  Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA  Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA  Bueno M.R.P., Camargo A.A., Carraro L.E.A., Carraro D.M., Carer H.,
RA  Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA  Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA  Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA  Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA  Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA  Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA  Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA  Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA  Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA  Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA  Menck C.F.M., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,
RA  Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA  Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA  de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA  Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA  Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA  de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA  da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA  da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA  de Souza A.P., Terezzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA  Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA  Zago M.A., Zatz M., Meldanis J., Setubal J.C.;
RT  "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL  Nature 406:151-159(2000).
DR  EMBL; AE004018; AAF84790.1; -.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 39 AA; 4501 MW; E085D64BE286D612 CRC64;

Query Match 15.3%; Score 33; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 1.2e+03;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 TFSNTSSTDICRPHQ 21
DB 24 SLKVTVSADVMRAHR 39

RESULT 14
Q9UD12 PRELIMINARY; PRT; 33 AA.
AC Q9UD12;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE AMGX PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95322983; PubMed=7599636;
RA Lench N.J., Winter G.B.;
RT "Characterisation of molecular defects in X-linked amelogenesis
RT imperfecta (AIH1).";
RL Hum. Mutat. 5:251-259(1995).
SQ SEQUENCE 33 AA; 3681 MW; D131F784BD7D8C93 CRC64;

Query Match 14.8%; Score 32; DB 4; Length 33;
Best Local Similarity 26.9%; Pred. No. 1.4e+03;
Matches 7; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 PCAPGTFSTSTSDICRPHQICNVV 26
```

```

Db 8 PCSPCRSHLCLRCSPCSPCLPCLFI 33

RESULT 15
Q9QL92 PRELIMINARY; PRT; 34 AA.
AC Q9QL92;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE 7.7 KDA PROTEIN.
OS Human adenovirus type 7a.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=85755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-1038;
RA Inada T., Mukoyama A., Yamadera S., Hashido M., Inoue S.;
RT "Epidemiology and genomic analysis of hexon, fiber and E3 region genes
RT of adenovirus type 7 in Japan.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104382; AAF14124.1; -.
SQ SEQUENCE 34 AA; 3939 MW; DF0B17CA2DA99972 CRC64;

Query Match 14.8%; Score 32; DB 12; Length 34;
Best Local Similarity 38.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 10 TTSSTDICRPHQI 22
DB 16 TTLNHDMAKPHYL 28

Search completed: February 12, 2002, 12:59:11
Job time: 181 sec
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:54:04 ; Search time 23.57 Seconds  
(without alignments)  
122.565 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_163\_201

Perfect score: 216

Sequence: 1 PCAPGTFNSTTSDICRPHQICNVVAIPGNASMDVCT 39

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 262775

Minimum DB seq length: 0

Maximum DB seq length: 39

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	21.3	17	20	AAW95323 Constant and variab
2	46	21.3	38	20	AAW12433 Human 5' EST secre
3	42	19.4	36	21	AAW37394 Human secreted pep
4	41	19.0	17	21	AAW51961 P. pastoris lysyl
5	41	19.0	17	21	AAW51973 P. pastoris lysyl
6	40	18.5	28	21	AAW64941 Human 5' EST relat
7	39	18.1	25	22	AAW16871 Peptide #3305 enco
8	39	18.1	25	22	AAW29355 Peptide #3392 enco
9	39	18.1	25	22	AAW04573 Peptide #3255 enco
10	39	18.1	26	20	AAW24434 Adenovirus hexon p
11	37.5	17.4	34	21	AAW44864 Human secreted pro

12	37	17.1	22	21	AAW51542 Yada homologous pe
13	37	17.1	23	20	AAW73416 Human secreted pro
14	37	17.1	27	22	AAW19434 Peptide #5968 enco
15	37	17.1	27	22	AAW32357 Peptide #6394 enco
16	37	17.1	37	22	AAW19721 Peptide #6155 enco
17	37	17.1	37	22	AAW32991 Peptide #7028 enco
18	36	16.7	22	16	AAW70346 NF-AT transcrip
19	36	16.7	22	21	AAW96558 Human NF-ATc1 anti
20	36	16.7	22	22	AAW66490 Human NF-AT peptid
21	36	16.7	23	19	AAW65481 Hepatitis B surfac
22	36	16.7	23	20	AAW97502 Antigenic site of
23	36	16.7	27	22	AAW80444 Gene #21 associate
24	36	16.7	28	14	AAW37510 I(21Cys, 22-32Pept
25	36	16.7	32	19	AAW65475 Hepatitis B surfac
26	36	16.7	32	19	AAW65476 Hepatitis B surfac
27	36	16.7	32	19	AAW65479 Hepatitis B surfac
28	36	16.7	32	19	AAW65480 Hepatitis B surfac
29	36	16.7	35	22	AAW28490 Peptide #2527 enco
30	36	16.7	38	20	AAW02252 A F-box protein se
31	36	16.7	38	22	AAW19313 Peptide #5747 enco
32	36	16.7	38	22	AAW32087 Peptide #6124 enco
33	35.5	16.4	17	11	AAW06092 Immunoreactive pep
34	35.5	16.4	35	22	AAW35518 Peptide #9555 enco
35	35.5	16.4	37	21	AAW91614 Human secreted pro
36	35	16.2	23	21	AAW27643 Human secreted pro
37	35	16.2	25	19	AAW44915 Spacer peptide for
38	35	16.2	27	17	AAW04374 Mu-conotoxin precu
39	35	16.2	31	6	AAW50156 Sequence of synthe
40	35	16.2	35	19	AAW44920 Spacer peptide for
41	35	16.2	36	18	AAW01811 Human papillomavir
42	35	16.2	39	22	AAW16939 Peptide #3373 enco
43	35	16.2	39	22	AAW29424 Peptide #3461 enco
44	35	16.2	39	22	AAW04632 Peptide #3314 enco
45	34.5	16.0	17	22	AAU05310 R1 and R2 peptide

#### ALIGNMENTS

#### RESULT 1

AAW95323

ID AAW95323 standard; Protein; 17 AA.

XX AC AAW95323;

XX AC AAW95323;

DT 15-MAR-1999 (first entry)

DE Costant and variable domain sequence of C. psittaci CPS92-106.

XX Chlamydia; cryptic phase; elementary body phase; replicating; probenidicid;  
KW antiporphyrin acid; immune response; infection; diagnostic; assay; MOMP;  
KW major outer membrane protein; autoImmune; inflammatory; porphyria;  
KW Ebstein Barr virus; antioxidant.

XX Chlamydia psittaci.

XX WO9850074-A2.

XX 12-NOV-1998.

XX 06-MAY-1998; 98WO-US09237.

XX 18-FEB-1998; 98US-0025521.

XX 06-MAY-1997; 97US-0045689.

XX 06-MAY-1997; 97US-0045739.

XX 06-MAY-1997; 97US-0045779.

XX 06-MAY-1997; 97US-0045780.

XX 06-MAY-1997; 97US-0045784.

XX 06-MAY-1997; 97US-0045787.

XX 14-AUG-1997; 97US-0911593.

XX 18-FEB-1998; 98US-0025174.

XX 18-FEB-1998; 98US-0025176.

PA (UYVA-) UNTV VANDERBILT.  
 XX  
 PI Mitchell WM, Stratton CW;  
 PR  
 DR WPI; 1999-059653/05.  
 XX  
 XX Composition with two agents effective against different stages of  
 PT chlamydial life cycle - comprises agent targetted against cryptic  
 PT phase, against elementary body phase, against replicating phase,  
 PT prohenicid and antiporphyric  
 XX  
 PS Claim 4; Fig 3; 138pp; English.  
 XX  
 CC The invention relates to the diagnosis and management of infections by  
 CC Chlamydia species. The invention provides a composition that comprises  
 CC at least two agents, where each of the agents is effective against a  
 CC different phase of the chlamydial life cycle. The agents are selected  
 CC from: (a) agents targetted against cryptic phase of chlamydial life  
 CC cycle; (b) agents targetted against elementary body phase of chlamydial  
 CC life cycle; (c) agents targetted against replicating phase of chlamydial  
 CC life cycle; (d) prohenicid, and (e) antiporphyric acid. The composition  
 CC is used to elicit a protective immune response to Chlamydia infection in  
 CC an animal or human and is applied until the animal or human tests  
 CC negative for Chlamydia infection. It is also used to treat biological  
 CC material infected with Chlamydia. Diagnostic kits for antibody assays  
 CC against recombinant major outer membrane protein (MOMP), and for DNA  
 CC amplification assays for chlamydial genes, are used to diagnose disease,  
 CC e.g. autoimmune disease, an inflammatory disease or a disease that  
 CC occurs in an immuno-compromised individual, associated with Chlamydia  
 CC infection. The kits are used to detect chlamydial elementary bodies in a  
 CC sample. They are also used to monitor and/or modify the course of therapy  
 CC in a patient. The treatment reduces the acellular load of infectious  
 CC Ebsstein Barr virus. The method is also used to treat porphyria, by  
 CC reducing the number of elementary bodies and applying a drug, e.g.  
 CC cimetidine, and antioxidants, to reduce the adverse effects associated  
 CC with porphyria. Sequences AA95320 to AA95323 represent constant and  
 CC variable domain sequences of various Chlamydia species.  
 XX  
 SQ Sequence 17 AA;

Query Match 21.3%; Score 46; DB 20; Length 17;  
 Best Local Similarity 64.3%; Pred. No. 12;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 CAPGTFSTSSSTD 15  
 Db 1 casgtsasnttvaad 14  
 || || || || || : | |  
 1 casgtsasnttvaad 14

RESULT 2  
 AAY12433  
 ID AAY12433 standard; Protein; 38 AA.  
 XX  
 AC AAY12433;  
 XX  
 DT 17-JUN-1999 (first entry)  
 XX  
 DE Human 5' EST secreted protein SEQ ID NO:464.  
 XX

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemotactic; chemokine; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition.  
 OS Homo sapiens.  
 XX  
 PN WO9906548-A2.  
 XX  
 PD 11-FEB-1999.  
 PR  
 XX

PF 31-JUL-1998; 98WO-IB01222.  
 XX  
 PR 01-AUG-1997; 97US-0905135.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 XX Duclert A, Dumas Milne Edwards J, Lacroix B;  
 PI  
 XX WPI; 1999-153778/13.  
 DR N-PSDB; AAX41266.  
 DR  
 XX

XX New nucleic acids encoding human secreted proteins - obtained from  
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,  
 PT kidney, lung, umbilical cord, placenta and colon tissue  
 XX  
 PS Claim 27; Page 768-769; 824pp; English.  
 XX

CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAY12261 to  
 CC AAY12514, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 38 AA;

Query Match 21.3%; Score 46; DB 20; Length 38;  
 Best Local Similarity 40.9%; Pred. No. 30;  
 Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 CAPGTFSTSSSTDICRPHQIC 23  
 Db 10 crpatldqatratpcrlsggc 31  
 | | | | | : | | | |  
 10 crpatldqatratpcrlsggc 31

RESULT 3  
 AAB37394  
 ID AAB37394 standard; Peptide; 36 AA.  
 XX  
 AC AAB37394;  
 XX  
 DT 20-FEB-2001 (first entry)  
 XX  
 DE Human secreted peptide #32 encoded by cDNA #47.  
 XX

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 OS Homo sapiens.  
 XX  
 PN WO200058335-A1.  
 XX  
 PD 05-OCT-2000.  
 XX  
 XX 22-MAR-2000; 2000WO-US07534.  
 XX  
 PR 26-MAR-1999; 99US-0126598.  
 PR 22-DEC-1999; 99US-0171504.  
 PR



CC peroxidases, lysyl oxidases, protein disulfide reductases, tyrosine  
 CC oxidases or sulphydryl oxidases is used to formulate compositions  
 CC containing active ingredients. (I) is useful for crosslinking protein  
 CC layers surrounding active ingredients in food, animal feed and  
 CC pharmaceutical products. The compositions can be formulated without  
 CC using chemical crosslinking agents. AAY51963-Y51974 represent fragments  
 CC of the Pichia pastoris lysyl oxidase protein which is used to  
 CC illustrate the method of the invention.

XX  
 XX  
 SQ Sequence 17 AA;

Query Match 19.0%; Score 41; DB 21; Length 17;  
 Best Local Similarity 70.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNT 10  
 ||||| ||  
 Db 7 pcpgvvvnt 16

RESULT 6  
 AAY64941  
 ID AAY64941 standard; Protein; 28 AA.

XX  
 AC AAY64941;

XX  
 DT 01-FEB-2000 (first entry)

XX  
 DE Human 5' EST related polypeptide SEQ ID NO:1102.

XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;  
 KW gene therapy; chromosome mapping; upstream regulatory sequence;  
 KW forensic; location; development; protein synthesis; stability;  
 KW regulation; identification.

OS Homo sapiens.

XX  
 PN WO9953051-A2.

XX  
 PD 21-OCT-1999.

XX  
 PF 09-APR-1999; 99WO-IB00712.

XX  
 PR 09-APR-1998; 98US-0057719.

XX  
 PR 28-APR-1998; 98US-0069047.

XX  
 PA (GEST ) GENSET.

XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX  
 DR WPI; 2000-038446/03.

XX  
 DR N-PSDB; AA242555.

XX  
 PT Novel secreted protein 5' expressed sequence tag sequences used in  
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures

XX  
 PS Claim 3; Page 687; 837pp; English.

XX  
 CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)  
 CC sequences, corresponding to human secreted proteins. AAY64651 to  
 CC AAY65438 represent the EST-related proteins corresponding to AA242265 to  
 CC AA243052. The 5' ESTs can be used for producing secreted human gene  
 CC products. They can be used to identify and isolate 5' untranslated  
 CC regions (UTRs) and upstream regulatory regions which control the  
 CC location, development stage, rate, and quantity of protein synthesis, as  
 CC well as stability of mRNA. The ESTs are also useful as probes for  
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can  
 CC also be used in forensic procedures to identify individuals, or in  
 CC diagnostic procedures to identify individuals having genetic diseases  
 CC resulting from abnormal gene expression. The products may also be used in  
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be  
 CC used for directing extracellular secretion of a polypeptide or the

CC insertion of a polypeptide into a membrane, or importing a polypeptide  
 CC into a cell. The proteins encoded by the EST sequences may be useful in  
 CC treating a variety of human conditions. Secreted proteins have  
 CC therapeutic value, and the identification of new secreted proteins is  
 CC valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent  
 CC sequences used in the exemplification of the present invention.

XX  
 SQ Sequence 28 AA;

Query Match 18.5%; Score 40; DB 21; Length 28;  
 Best Local Similarity 41.2%; Pred. No. 1.4e+02;  
 Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAPGTFSTNTSSSTDICR 18  
 | | | | | | | | | |  
 Db 10 crpatldqatratpccr 26

RESULT 7

AAM16871

ID AAM16871 standard; Protein; 25 AA.

XX  
 AC AAM16871;

XX  
 DT 12-OCT-2001 (first entry)

XX  
 DE Peptide #3305 encoded by probe for measuring cervical gene expression.

XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.

OS Homo sapiens.

XX  
 PN WO200157278-A2.

XX  
 PD 09-AUG-2001.

XX  
 PF 30-JAN-2001; 2001WO-US00670.

XX  
 PR 04-FEB-2000; 2000US-0180312.

XX  
 PR 26-MAY-2000; 2000US-0207456.

XX  
 PR 30-JUN-2000; 2000US-0608408.

XX  
 PR 03-AUG-2000; 2000US-0632366.

XX  
 PR 21-SEP-2000; 2000US-0234687.

XX  
 PR 27-SEP-2000; 2000US-0236359.

XX  
 PR 04-OCT-2000; 2000GB-0024263.

XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
 DR WPI; 2001-488901/53.

XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -

XX  
 PS Claim 27; SEQ ID No 21697; 487pp; English.

XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.

XX  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 25 AA;

Query Match 18.1%; Score 39; DB 22; Length 25;  
 Best Local Similarity 45.0%; Pred. No. 1.6e+02;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNTSSDTCRPH 20  
 | | | | | | | | | |  
 Db 3 pwfeglpshsttttsimrrh 22

## RESULT 8

AAM29355  
 ID AAM29355 standard; Protein; 25 AA.

AC AAM29355;

DT 17-OCT-2001 (first entry)

DE Peptide #3392 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 29624; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

XX Sequence 25 AA;

Query Match 18.1%; Score 39; DB 22; Length 25;  
 Best Local Similarity 45.0%; Pred. No. 1.6e+02;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNTSSDTCRPH 20  
 | | | | | | | | | |  
 Db 3 pwfeglpshsttttsimrrh 22

## RESULT 9

AAM04573  
 ID AAM04573 standard; Protein; 25 AA.

XX

AC AAM04573;

DT 09-OCT-2001 (first entry)

DE Peptide #3255 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

OS Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US00661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression  
 in a human breast -

PS Claim 27; SEQ ID No 13313; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes  
 (see AAI00010-AAI10067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in  
 CC a human breast sample, where the probe hybridises at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 25 AA;

Query Match 18.1%; Score 39; DB 22; Length 25;  
 Best Local Similarity 45.0%; Pred. No. 1.6e+02;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 PCAPGTFSTNTSSDTCRPH 20  
 | | | | | | | | | |  
 Db 3 pwfeglpshsttttsimrrh 22

## RESULT 10

AAI24434  
 ID AAY24434 standard; peptide; 26 AA.

XX AAY24434;

DT 27-SEP-1999 (first entry)

DE Adenovirus hexon protein heterologous ligand #2.

XX Adenovirus; RGD sequence; adenoviral vector; capsid protein; ligand;



```

XX AAB51542;
AC
XX
XX 15-FEB-2001 (first entry)
DT
XX
XX Yada homologous peptide #5.
DE
XX
XX proteobacteria; extracellular domain; virulence determinant; Yada;
KW adhesin; proteobacterial infection prevention; vaccine.
KW
XX
XX Thiobacillus ferrooxidans.
OS
XX
XX WO200061165-A1.
PN
XX
XX 19-OCT-2000.
PD
XX
XX 13-APR-2000; 2000WO-US09866.
PF
XX
XX 13-APR-1999; 99US-0129073.
PR
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA
XX (SMIK ) SMITHKLINE BEECHAM PLC.
PA
XX
XX Lupas AN;
PI
XX
XX WPI: 2000-647397/62.
DR
XX
XX An isolated polypeptide conserved in proteobacterial extracellular
PT domains used in the treatment and prevention of bacterial infections -
PT
XX
XX Example 5; Page 59; 85pp; English.
PS
XX
XX This invention relates to peptides AAB51512 - AAB51537 which represent
CC conserved proteobacterial extracellular domains. Sequences
CC AAB51538 - AAB51618 represent peptides homologous to Yada, a Yersinia
CC adhesin which is an important virulence determinant of the Yersinia
CC species. The invention includes an antibody which binds to the
CC proteobacterial extracellular peptides, and an immunogenic composition
CC containing the antibody used as a vaccine to prevent infection by a
CC proteobacteria. The polypeptides and antibodies are useful in the
CC treatment and prevention of proteobacterial infections. The polypeptides
CC can also be used to identify compounds which antagonize the binding of a
CC bacterial adhesin to its ligand. The host cell can be used to produce
CC the polypeptides in a suitable culture system. The composition can be
CC used to vaccinate a patient against a proteobacterial infection.
XX
XX Sequence 22 AA;
SQ
Query Match 17.1%; Score 37; DB 21; Length 22;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
QY 3 APGTFSTNTSSDTCRPHQI 22
Db ||||| | |||
5 apgtlsgt--stdavngsqi 22
RESULT 13
AAW73416
ID AAW73416 standard; Protein: 23 AA.
XX
XX AAW73416;
XX
XX 19-FEB-1999 (first entry)
DT
XX
XX Human secreted protein encoded by Gene No. 20.
DE
XX
XX Secreted protein; human; protein therapy; gene therapy; blood disorder;
KW pathological condition; diagnosis; cancer; neurological disorder;
KW developmental abnormality; foetal deficiency; leukaemia; hepatic disease;
KW immune system disorder; Alzheimer's disease; cognitive disorder;
KW schizophrenia; prostate disease; autoimmune disorder; AIDS.

```

```

XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 23
FT /note= "unspecified amino acid"
XX
XX WO9854206-A1.
PN
XX
XX 03-DEC-1998.
PD
XX
XX 28-MAY-1998; 98WO-US10868.
PF
XX
XX 29-AUG-1997; 97US-0056296.
PR
XX 30-MAY-1997; 97US-0044039.
PR
XX 30-MAY-1997; 97US-0048093.
PR
XX 30-MAY-1997; 97US-0048101.
PR
XX 30-MAY-1997; 97US-0048190.
PR
XX 30-MAY-1997; 97US-0048356.
PR
XX 30-MAY-1997; 97US-0050935.
PR
XX 29-AUG-1997; 97US-0056250.
PR
XX 29-AUG-1997; 97US-0056293.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Carter KC, Dillon PJ, Endress GA, Feng P, NI J;
PI Rosen CA, Ruben SM, Yu G;
XX
XX WPI: 1999-070209/06.
DR
XX N-PSDB; AAV08830.
DR
XX
XX New isolated human genes - useful for diagnosis and treatment of,
PT e.g. cancers, neurological disorders, immune diseases, developmental
PT disorders or blood disorders
XX
XX Claim 11; Page 157; 188pp; English.
XX
XX This sequence is encoded by a cDNA of the invention, designated
CC Gene No. 20. This sequence represents a human secreted protein, and is
CC expressed ubiquitously, including T-cells and amygdala.
CC The DNA sequences of the invention and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the DNA sequences.
CC Specific uses are described for each of the DNA sequences and the encoded
CC proteins, based on which tissues they are most highly expressed in, and
CC include developing products for the diagnosis or treatment of cancer,
CC tumours, neurological disorders, developmental abnormalities and foetal
CC deficiencies, blood disorders, leukaemias, diseases of the immune system
CC (including allergies or asthma), hepatic disease, Alzheimer's and
CC cognitive disorders, schizophrenia, prostate diseases, autoimmune
CC disorders and AIDS. The polypeptides are also useful for identifying
CC their binding partners.
XX
XX Sequence 23 AA;
SQ
Query Match 17.1%; Score 37; DB 20; Length 23;
Best Local Similarity 43.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 2 CAPGTFSTNTSSDTC 17
Db ||| | |||
5 cpggaagtacssac 20
RESULT 14
AAW19434
ID AAW19434 standard; Protein: 27 AA.
XX
XX AAW19434;
XX
XX

```

DT 12-OCT-2001 (first entry)  
XX Peptide #5868 encoded by probe for measuring cervical gene expression.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX Homo sapiens.  
OS  
XX WO200157278-A2.  
PN 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US006670.  
PF 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
PI Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human cervical epithelial cells -  
PT Claim 27; SEQ ID No 24260; 487pp; English.  
PS The present invention relates to human single exon nucleic acid probes  
XX (SENP; see AAI10068-AAI28459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 27 AA;  
  
Query Match 17.1%; Score 37; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 19 PHQICN 24  
Db 13 phqlcn 18  
  
RESULT 15  
AAM32357  
ID AAM32357 standard; Protein; 27 AA.  
XX  
AC AAM32357;  
XX  
DT 17-OCT-2001 (first entry)  
XX Peptide #6394 encoded by probe for measuring placental gene expression.  
DE Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX Homo sapiens.  
OS

PN WO200157272-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US006663.  
PF 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
PI Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human placenta -  
PT Claim 27; SEQ ID No 32626; 654pp; English.  
PS The present invention relates to single exon nucleic acid probes (SENP;  
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders.  
SQ Sequence 27 AA;  
  
Query Match 17.1%; Score 37; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 19 PHQICN 24  
Db 13 phqlcn 18  
  
Search completed: February 12, 2002, 12:56:04  
Job time: 120 sec





Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	216	100.0	39	1	US-08-050-319B-41	Sequence 41, Appl
2	216	100.0	39	2	US-08-465-982-41	Sequence 41, Appl
3	77	35.6	15	6	5395760-10	Patent No. 5395760
4	72	33.3	15	1	US-08-221-583-45	Sequence 45, Appl
5	72	33.3	15	1	US-08-221-583-45	Sequence 45, Appl
6	72	33.3	15	5	PCr-US95-04018-46	Sequence 46, Appl
7	72	33.3	15	5	PCr-US95-04018-46	Sequence 45, Appl
8	71	32.9	15	1	US-08-221-583-44	Sequence 46, Appl
9	71	32.9	15	5	PCr-US95-04018-44	Sequence 44, Appl
10	70	32.4	15	5	US-08-221-583-49	Sequence 44, Appl
11	70	32.4	15	5	PCr-US95-04018-49	Sequence 49, Appl
12	67	31.0	15	1	US-08-221-583-47	Sequence 49, Appl
13	67	31.0	15	5	PCr-US95-04018-47	Sequence 47, Appl
14	66	30.6	15	1	US-08-221-583-48	Sequence 47, Appl
15	66	30.6	15	5	PCr-US95-04018-48	Sequence 48, Appl
16	57	26.4	15	1	US-08-221-583-50	Sequence 48, Appl
17	57	26.4	15	5	PCr-US95-04018-50	Sequence 50, Appl
18	48	22.2	20	2	US-08-126-016-24	Sequence 50, Appl
19	45	20.8	15	1	US-08-221-583-51	Sequence 24, Appl
20	45	20.8	15	5	PCr-US95-04018-51	Sequence 51, Appl
21	39	18.1	23	3	US-09-101-146-58	Sequence 51, Appl
22	37.5	17.4	20	2	US-08-126-016-25	Sequence 58, Appl
23	36	16.7	22	2	US-08-124-981A-28	Sequence 25, Appl
24	36	16.7	22	3	US-09-037-190-31	Sequence 28, Appl
25	36	16.7	22	3	US-09-037-192-31	Sequence 31, Appl
26	36	16.7	22	4	US-09-037-143-31	Sequence 31, Appl
27	36	16.7	22	4	US-09-049-691-31	Sequence 31, Appl

```
RESULT 2
US-08-465-982-41
; Sequence 41, Application US/08465982
; Patent No. 5863786
; GENERAL INFORMATION:
; APPLICANT: M. Feldmann, P.W. Gray,
; APPLICANT: M.J.C. Turner, F.M. Brennan
; TITLE OF INVENTION: Modified human TNFalpha (Tumor
; TITLE OF INVENTION: Necrosis Factor alpha) Receptor
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robbins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,982
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,319
; FILING DATE: 10-May-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Robbins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5150-0030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-982-41

Query Match 100.0%; Score 216; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.9e-22;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCAPGTFSTSTDCRPHQICNVVAIPGNASMDVCT 39
Db 1 PCAPGTFSTSTDCRPHQICNVVAIPGNASMDVCT 39

RESULT 3
5395760-10
; Patent No. 5395760
; APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN,
; M. PATRICIA
; TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
; B-RECEPTORS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/523,635
; FILING DATE: 10-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 421,417
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: 405,370
; FILING DATE: 11-SEP-1989
```

```
; APPLICATION NUMBER: 403,241
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 10
; LENGTH: 15
5395760-10

Query Match 35.6%; Score 77; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 ICNVVAIPGNASMDA 36
Db 1 ICNVVAIPGNASMDA 15

RESULT 4
US-08-221-583-45
; Sequence 45, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,583
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-221-583-45

Query Match 33.3%; Score 72; DB 1; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0013;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTFSTSTSTDCR 19
Db 1 GTFSTSTSTDIAP 15

RESULT 5
US-08-221-583-46
; Sequence 46, Application US/08221583
; Patent No. 5486595
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 62
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
;; STREET: One Liberty Place 46th Floor  
;; CITY: Philadelphia  
;; STATE: Pennsylvania  
;; COUNTRY: USA  
;; ZIP: 19403  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/221,583  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Deluca, Mark  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: CCOR-0185  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 46:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-221-583-46

Query Match 33.3%; Score 72; DB 1; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.0013;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 SNTSSTDICRPHQI 22  
Db 1 SNTSSTDICRPHQI 15

RESULT 6  
PCT-US95-04018-45  
;; Sequence 45, Application PC/TUS9504018  
;; GENERAL INFORMATION:  
;; APPLICANT: Heavner, George A.  
;; APPLICANT: Kruszynski, Marian  
;; APPLICANT: Mervic, Miljenko  
;; APPLICANT: Weber, Robert W.  
;; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
;; NUMBER OF SEQUENCES: 76  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
;; ADDRESSEE: Norris  
;; STREET: One Liberty Place 46th Floor  
;; CITY: Philadelphia  
;; STATE: Pennsylvania  
;; COUNTRY: USA  
;; ZIP: 19403  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WordPerfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/04018  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,580  
;; FILING DATE: 01-APR-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,583

;; FILING DATE: 01-APR-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,581  
;; FILING DATE: 01-APR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Deluca, Mark  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: CCOR-0232  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 45:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
PCT-US95-04018-45

Query Match 33.3%; Score 72; DB 5; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.0013;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GTFSTSTSDICRP 19  
Db 1 GTFSTSTSDIARP 15

RESULT 7  
PCT-US95-04018-46  
;; Sequence 46, Application PC/TUS9504018  
;; GENERAL INFORMATION:  
;; APPLICANT: Heavner, George A.  
;; APPLICANT: Kruszynski, Marian  
;; APPLICANT: Mervic, Miljenko  
;; APPLICANT: Weber, Robert W.  
;; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
;; NUMBER OF SEQUENCES: 76  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
;; ADDRESSEE: Norris  
;; STREET: One Liberty Place 46th Floor  
;; CITY: Philadelphia  
;; STATE: Pennsylvania  
;; COUNTRY: USA  
;; ZIP: 19403  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WordPerfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/04018  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,580  
;; FILING DATE: 01-APR-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,583  
;; FILING DATE: 01-APR-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/221,581  
;; FILING DATE: 01-APR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Deluca, Mark  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: CCOR-0232  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 46:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-46

Query Match 33.3%; Score 72; DB 5; Length 15;  
Best Local Similarity 93.3%; Pred. No. 0.0013;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 SNTTSSTDICRPHQI 22  
| | | | | | | | | | | | | | |  
Db 1 SNTTSSTDICRPHQI 15

RESULT 8  
US-08-221-583-44  
; Sequence 44, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-44

Query Match 32.9%; Score 71; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTGFSNTTSSTDI 16  
| | | | | | | | | | | | | | |  
Db 2 APTGFSNTTSSTDI 15

RESULT 9  
PCT-US95-04018-44  
; Sequence 44, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko

; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-44

Query Match 32.9%; Score 71; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0018;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTGFSNTTSSTDI 16  
| | | | | | | | | | | | | | |  
Db 2 APTGFSNTTSSTDI 15

RESULT 10  
US-08-221-583-49  
; Sequence 49, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

;  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-49

Query Match 32.4%; Score 70; DB 1; Length 15;  
Best Local Similarity 92.9%; Pred. No. 0.0025;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 RPHQICNVVAIPGN 31  
||||| |||||||  
Db 2 RPHQIANVVAIPGN 15

## RESULT 11

PCT-US95-04018-49  
; Sequence 49, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian  
; APPLICANT: Mervic, Miljenko  
; APPLICANT: Weber, Robert W.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: Norris  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04018  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,580  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,583  
; FILING DATE: 01-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/221,581  
; FILING DATE: 01-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0232  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439

;  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-49

Query Match 32.4%; Score 70; DB 5; Length 15;  
Best Local Similarity 92.9%; Pred. No. 0.0025;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 RPHQICNVVAIPGN 31  
||||| |||||||  
Db 2 RPHQIANVVAIPGN 15

## RESULT 12

US-08-221-583-47  
; Sequence 47, Application US/08221583  
; Patent No. 5486595  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595r1s  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19403  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25:mdctcMod.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/221,583  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: CCOR-0185  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-221-583-47

Query Match 31.0%; Score 67; DB 1; Length 15;  
Best Local Similarity 86.7%; Pred. No. 0.0062;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQICNV 25  
||||| |||||||  
Db 1 TSSTDICRPHQIANV 15

## RESULT 13

PCT-US95-04018-47  
; Sequence 47, Application PC/TUS9504018  
; GENERAL INFORMATION:  
; APPLICANT: Heavner, George A.  
; APPLICANT: Kruszynski, Marian

APPLICANT: Mervic, Miljenko  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: Norris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04018  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,580  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,583  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,581  
FILING DATE: 01-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04018-47

Query Match 31.0%; Score 67; DB 5; Length 15;  
Best Local Similarity 86.7%; Pred. No. 0.0062;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 TSSTDICRPHQICNV 25  
||||| ||||| ||  
Db 1 TSSTDICRPHQIANV 15

RESULT 14  
US-08-221-583-48  
Sequence 48, Application US/08221583  
Patent No. 5486595  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5486595ris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25.mdctcMod.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/221,583  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0185  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-221-583-48

Query Match 30.6%; Score 66; DB 1; Length 15;  
Best Local Similarity 86.7%; Pred. No. 0.0084;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 TDIARPHQICNVVAVI 28  
||||| ||||| |||||  
Db 1 TDIARPHQIANVAVI 15

RESULT 15  
PCT-US95-04018-48  
Sequence 48, Application PC/TUS9504018  
GENERAL INFORMATION:  
APPLICANT: Heavner, George A.  
APPLICANT: Kruszynski, Marian  
APPLICANT: Mervic, Miljenko  
APPLICANT: Weber, Robert W.  
TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESSEE: Norris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04018  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,580  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,583  
FILING DATE: 01-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/221,581  
FILING DATE: 01-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CCOR-0232  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US95-04018-48

Query Match 30.6%; Score 66; DB 5; Length 15;  
Best Local Similarity 86.7%; Pred. No. 0.0084;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 14 TDICRPHQICNVVAI 28  
   | | | | | | | | | |  
Db 1 TDIARPHQIANVVAI 15

Search completed: February 12, 2002, 12:56:23  
Job time: 114 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 12:58:45 ; Search time 13.29 Seconds  
(without alignments)  
326.708 Million cell updates/sec

Title: US-09-800-909-2\_COPY\_201\_257  
Perfect score: 302  
Sequence: 1 TSTSPTRSMAPGAVHLPQPV.....STSFLLPMGPPSPPAEGSTGCD 57  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 13308

Minimum DB seq length: 0  
Maximum DB seq length: 57

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:★  
1: pir1:★  
2: pir2:★  
3: pir3:★  
4: pir4:★

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	23.5	54	2 A60726	mucin, gallbladder
2	56.5	18.7	52	2 S63324	hypothetical prote
3	55.5	18.4	42	2 I70082	glycoprotein Ib al
4	51	16.9	25	2 I40692	cenA protein (IgaI
5	50	16.6	46	2 S07073	arabinogalactan pr
6	47	15.6	31	2 T36022	small hypothetical
7	44.5	14.7	57	2 S10782	salivary protein P
8	44.5	14.7	57	2 S16587	hypothetical prote
9	44	14.6	28	2 I56139	MHC class I HLA-J
10	43.5	14.4	47	2 S32108	sepiapterin reduct
11	43.5	14.4	54	2 D81737	hypothetical prote
12	43	14.2	31	2 I54515	pre-B cell Ig lamb
13	43	14.2	39	2 A46662	collagen alpha 2(V
14	42	13.9	36	2 A37172	collagen alpha 1(X
15	42	13.9	48	2 I46522	troponin T 2fa - r
16	42	13.9	55	2 S08424	H+-transporting AT
17	42	13.9	55	2 T11184	H+-transporting AT
18	41	13.6	42	2 T07030	extensin - tomato
19	40.5	13.4	50	2 A29789	mucin - sheep (fra
20	40	13.2	29	2 G39690	neural cell adhesi
21	40	13.2	50	2 H64801	hypothetical prote
22	40	13.2	51	2 A33756	dorsal protein - f
23	40	13.2	51	2 G72801	gp18 protein - myc
24	40	13.2	52	2 S58216	hypothetical prote
25	40	13.2	54	2 T11131	ATP synthase subun
26	40	13.2	55	2 T11105	H+-transporting AT
27	40	13.2	57	2 I58120	gene insl protein
28	39.5	13.1	52	2 S01945	myosin catalytic I
29	39	12.9	27	2 S51176	aspartate transcar

30 39 12.9 53 2 S17672 hypothetical prote  
31 38.5 12.7 51 2 C81117 hypothetical prote  
32 38 12.6 53 2 S23202 kappa-casein - bov  
33 38 12.6 54 2 S14338 serpin I - horse (  
34 38 12.6 55 2 S29770 DNA-binding protei  
35 38 12.6 55 2 A37238 autoimmune epitope  
36 37.5 12.4 49 2 S25433 neural cell adhesi  
37 37.5 12.4 51 2 T29481 hypothetical prote  
38 37.5 12.4 57 2 E41715 hypothetical prote  
39 37.5 12.4 57 2 T36648 hypothetical prote  
40 37 12.3 30 2 JC1360 hypothetical 3K pr  
41 37 12.3 38 2 S68260 hypothetical prote  
42 37 12.3 45 2 S24713 Ig alpha chain, tr  
43 37 12.3 49 2 S72213 beta-fructofuranos  
44 37 12.3 51 2 B32040 dihydrolipoamide S  
45 37 12.3 55 2 A82795 hypothetical prote

ALIGNMENTS

RESULT 1  
A60726  
mucin, gallbladder - bovine (fragments)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 30-Sep-1993  
C:Accession: A60726  
R:Afdhal, N.H.; Offner, G.D.; Smith, B.F.  
Gastroenterology 99, 1493-1501, 1990  
A:Title: Characterization of bovine gallbladder mucin. Amino acid sequences of trypti  
A:Reference number: A60726; MUID:91007106  
A:Accession: A60726  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-54 <AFD>

Query Match 23.5%; Score 71; DB 2; Length 54;  
Best Local Similarity 35.2%; Pred. No. 2.3;  
Matches 19; Conservative 6; Mismatches 25; Indels 4; Gaps 2;

Qy 2 STSPTRSMAPGAVHLPQVSTRSQHT-OPTPEPSTAPSTSF---LLPMGPPSPA 51  
Db 1 TTTTPTVPLPGSPQTTTSSPLPETPTPTPTVPSPSPQDTEPSGPQPPS 54

RESULT 2  
S63324  
hypothetical protein YNL338w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein NO170  
C:Species: Saccharomyces cerevisiae  
C:Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 05-Nov-1999  
C:Accession: S63324  
R:Obermaier, B.; Piravandi, E.; Rinke, M.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63317  
A:Accession: S63324  
A:Molecule type: DNA  
A:Residues: 1-52 <OBE>  
A:Cross-references: EMBL:Z71614; NID:g1302466; PID:G239576; GSPDB:GN000  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YNL338W  
A:Map position: 14L

Query Match 18.7%; Score 56.5; DB 2; Length 52;  
Best Local Similarity 35.7%; Pred. No. 4.2;  
Matches 15; Conservative 4; Mismatches 20; Indels 3; Gaps 2;

Qy 8 SMAPGAVHLPQVSTRSQHTOPTPEPSTAPSTSFLLPMGPPSP 49  
Db 11 SMQYSDIYIPTPTPTHTHTT-PTPHPH--PTHHTHTHTNP 49



```

Plant Mol. Biol. 17, 155-156, 1991
A;Title: DNA sequence of a mitochondrial plasmid from Chenopodium album.
A;Reference number: S16587; MUID:91329724
A;Accession: S16587
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-57 <DOE>
A;Cross-references: EMBL:X58911

Query Match      14.7%; Score 44.5; DB 2; Length 57;
Best Local Similarity 34.2%; Pred. No. 5.3e+02;
Matches 13; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY      16 LPQPVSTRSHTQTPPESTAPSTSFLLPMGCPSPPAEG 53
      || : : | : || || : : | : | : |
DB      21 LPWYLRLTFRVP-PDPQAPNFNHSLSLNPPPSAAG 57

RESULT 9
I56139
MHC class I HLA-J antigen - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C;Accession: I56139; I72807
R;Messier, G.; Zemmour, J.; Orr, H.T.; Parham, P.; Weiss, E.H.; Girdlestone, J.

```

Query Match	14.7%	Score 44.5;	DB 2;	Length 57;
Best Local Similarity	34.3%;	Pred. No. 5.3e+02;		
Matches 13; Conservative	6;	Mismatches 18;	Indels 1;	Gaps 1;
Qy	16	LPQPVTRSOHTQTPPESTAPSTSFLLPMGSPPAEG	53	
Db	21	LPWYRLTLTRVPD-PDPSQAPNFNHLSLSLNPPSAAVG	57	

```

RESULT      9
I56139
MHC class I HLA-J antigen - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I56139; I72807

```

R; Messer, G.; Zemmour, J.; Orr, H.T.; Parham, P.; Weiss, E.H.; Girdlestone, J. *J. Immunol.* 148, 4043-4053, 1992

A:Name: HLA-A\*01:01, 40% <RES>  
A>Title: HLA-A\*, a second inactivated class I HLA gene related to HLA-G and HLA-A. Implied  
A:Reference number: I56139; MUID:92291530  
A:Accession: I56139  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-28 <RES>  
A:Cross-references: GB:M80468; NID:g188481; PIDN:AAA36306.1; PID:g188482  
A:Accession: I72807  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-28 <RES>  
A:Cross-references: GB:M80469; NID:g188483; PIDN:AAA36307.1; PID:g188484  
C:Genetics:  
A:Introns: 24/3  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match	14.6%	Score 44;	DB 2;	Length 28;
Best Local Similarity	40.3%;	Pred. No. 2.8+02;		
Matches	9;	Conservative	3;	Mismatches 10;
				Indels
				0;
				Gaps
				0;

Qy	17	QPVSTRSQHTOPTPEFSTAPS	38
		! ! ! ! ! ! ! ! ! !	
Db	6	PEPSSCCSRGMPWPWPGRAPT	27

```

RESULT 10
S321108
sepiapterin reductase (EC 1.1.1.153) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 12-Sep-1997
C:Accession: S321108
R;Maier, J.
submitted to the EMBL Data Library, March: 1993
A:Reference number: S321108
A:Accession: S321108
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-47 <MAI>
A:cross-references: EMBL:221947
C:Keywords: oxidoreductase

```

Query Match	14.4%	Score 43.5;	DB 2;	Length 47;
Best Local Similarity	39.3%	Pred. NO. 5.3e+02;		
Matches 11;	Conservative	3;	Mismatches 13;	Indels 1;
				Gaps 1;

R;Mayne, R.; Brexton, R.G.; Mayne, P.M.; Baker, J.R.

J. Biol. Chem. 288, 9381-9386, 1993

A;Title: Isolation and characterization of the chains of type V/type XI collagen present

A;Reference number: A46662; MUID:93252802

A;Accession: A46662

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-39 <MAY>

A;Experimental source: vitreous humor

A;Note: sequence extracted from NCBI backbone (NCBIP:131547)

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Search completed: February 12, 2002, 13:00:37

Job time: 112 sec

Query Match 14.2%; Score 43; DB 2; Length 39;

Best Local Similarity 53.3%; Pred. No. 4.8e+02;

Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 43 LPMGPPPPAEGSTGD 57

Db 14 LPGPPPPGEAGPGD 28

#### RESULT 14

A37172

collagen alpha 1(XII) chain-like, skin and tendon - bovine (fragments)

C;Species: Bos primigenius taurus (cattle)

C;Date: 31-Jan-1992 #sequence\_revision 31-Jan-1992 #text\_change 19-Oct-1995

C;Accession: A37172

R;Dublet, B.; Van Der Rest, M.

Ann. N. Y. Acad. Sci. 580, 436-439, 1989

A;Title: Comparison between chicken type XII collagen and bovine homologues.

A;Reference number: A37172

A;Accession: A37172

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-36 <DUB>

Query Match 13.9%; Score 42; DB 2; Length 36;

Best Local Similarity 41.4%; Pred. No. 5.4e+02;

Matches 12; Conservative 3; Mismatches 8; Indels 6; Gaps 3;

QY 25 QHTQPTPEPSTA-PSTSFLLPMGPPPAE 52

Db 3 QERSP-PQANAVPSX---PASPSPLQ 26

#### RESULT 15

I46522

troponin T 2fa - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 13-Aug-1999

C;Accession: I46522

R;Briggs, M.M.; Lin, J.J.; Schachat, F.H.

J. Muscle Res. Cell. Motil. 8, 1-12, 1987

A;Title: The extent of amino-terminal heterogeneity in rabbit fast skeletal muscle tropo

A;Reference number: I46522; MUID:87251333

A;Accession: I46522

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-48 <BRI>

A;Cross-references: EMBL:U04975; NID:g440810; PIDN:AAA16028.1; PID:g440811

C;Superfamily: troponin T

Query Match 13.9%; Score 42; DB 2; Length 48;

Best Local Similarity 37.9%; Pred. No. 7.3e+02;

Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 2;

QY 10: APCAHLPPQVTRSQHTQTPPS-TAP 37

Db 25 SPAEVHEPEV-----HEEEKPRKLTAP 48



OC homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]

```

DR   InterPro: IPR001421; ATP-synt_8.
DR   Pfam: PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT   TRANSMEM 4      24      POTENTIAL.
SQ   SEQUENCE 55 AA; 6481 MW; E85C81E63DB48B15 CRC64;

Query Match 13.9%; Score 42; DB 1; Length 55;
Best Local Similarity 32.0%; Pred. No. 5.5e+02;
Matches 8; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 16 LPPQVSTRSQHTQTPPEPSTAPSTS 40
    ||| :||: |||:
DB 24 LPPKVMHAHTFPNEPSPQGMTPKTA 48

RESULT 4
VG18_BPMD2 STANDARD; PRT; 51 AA.
ID AC 064211;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 18 PROTEIN (GP18).
DN 18.
GN Mycobacteriophage D29.
OS Mycobacteriophage D29.
OC Viruses.
OX NCBI_TaxID=283369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage
   evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF022214; AAC18458.1;
DR SEQUENCE 51 AA; 5680 MW; EF85B1AFF5786A34 CRC64;

Query Match 13.2%; Score 40; DB 1; Length 51;
Best Local Similarity 33.3%; Pred. No. 7.5e+02;
Matches 11; Conservative 4; Mismatches 14; Indels 4; Gaps 1;

QY 7 RSMAPGAVHLPPQVSTRSQHTQTPPEPSTAPST 39
    ||| :||: |||: |||:
DB 21 RLIAAGGWRKPRKPRPT----TKPKPAPKQEPAT 49

RESULT 5
ATP8_FELSU STANDARD; PRT; 55 AA.
ID AC 079674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
DN MTATP8 OR ATP8.
GN Pelomedusa subrufa (African side-necked turtle).
OS Mitochondrion.
OG Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Pleurodira; Pelomedusidae; Pelomedusa.
OX NCBI_TaxID=44522;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Zardoya R.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
CC (CFO) (SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.  
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF039066; AAD05054.1; -.  
DR InterPro; IPR001421; ATP-synt\_8.  
DR Pfam; PF00895; ATP-synt\_8; 1.  
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
FT TRANSMEM 4 24 POTENTIAL.  
SQ SEQUENCE 55 AA; 6536 MW; D8D4BCF8F651A001 CRC64;  
-----  
Query Match 13.2%; Score 40; DB 1; Length 55;  
Best Local Similarity 36.0%; Pred. No. 8.1e+02;  
Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
-----  
QY 15 HLPQVSTRSQHTQPTPEPSTAPST 39  
|:| | :| | | | |  
Db 31 HIPNNSPTKKNKLLTTPMPWPWT 55  
-----  
RESULT 6  
TAGP\_HUMAN STANDARD; PRT; 56 AA.  
ID TAGP\_HUMAN  
AC Q9Y3F1;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TAP2-ASSOCIATED 6.5 KDA POLYPEPTIDE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li H.C.;  
RT "A 56 aa polypeptide with phosphorylation motif, potentially  
RT associated with Tap2 isoform activity."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MAY BE ASSOCIATED WITH TAP2 ISOFORM ACTIVITY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF152583; AAD32715.1; -.  
DR SEQUENCE 56 AA; 6535 MW; ACD5D223E2C3BA CRC64;  
-----  
Query Match 12.9%; Score 39; DB 1; Length 56;  
Best Local Similarity 40.0%; Pred. No. 1e+03;  
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
-----  
QY 30 TPEPSTAPSTSFLLPMGPSP 49  
|:| | :|:| | | | |  
Db 6 TPQILTSFVSILSLPSP 25

RESULT 7  
MLEV\_MOUSE  
ID MLEV\_MOUSE STANDARD; PRT; 51 AA.  
AC P09542;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MYOSIN LIGHT CHAIN 1, SLOW-TWITCH MUSCLE B/VENTRICULAR ISOFORM  
DE (FRAGMENT).  
GN MYL3 OR MLC1V.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H; TISSUE=Spleen;  
RX MEDLINE=89057447; PubMed=3194193;  
RA Cohen A., Barton P.J.R., Robert B., Garner I., Alonso S.,  
RA Buckingham M.E.;  
RT "Promoter analysis of myosin alkali light chain genes expressed in  
RT mouse striated muscle."  
RL Nucleic Acids Res. 16:10037-10052(1988).  
CC -!- SUBUNIT: MYOSIN IS A HEXAMER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.  
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS  
CC PROTEIN DOES NOT BIND CALCIUM.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X12972; CAA31415.1; -.  
DR PIR; S01945; S01945.  
DR HSSP; P04002; IATF.  
KW Myosin; Muscle protein; Multigene family.  
FT INTR\_MET 0  
FT NON\_TER 51 51  
SQ SEQUENCE 51 AA; 5085 MW; CE513ECBA3C8258D CRC64;  
-----  
Query Match 12.6%; Score 38; DB 1; Length 51;  
Best Local Similarity 26.3%; Pred. No. 1.1e+03;  
Matches 10; Conservative 6; Mismatches 20; Indels 2; Gaps 1;  
-----  
QY 17 POPVSTRSQHTQPT--TPEPSTAPSTSFLLPMGPSPPAE 52  
|:| | :| | | | | :| | | | |  
Db 5 PEPKDDAKAAAPKAPAPAAAPAAAPAAAPAAPEPKE 42  
-----  
RESULT 8  
ATP8\_PAROL STANDARD; PRT; 55 AA.  
ID ATP8\_PAROL  
AC Q9N9D5;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).  
GN MTATP8 OR ATP8.  
OS Paralichthys olivaceus (Flounder).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidei; Bothidae; Paralichthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Saitoh K., Hayashizaki K., Yokoyama Y., Asahida T., Toyohara H.,  
RA Yamashita Y.;





OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
OX Rhodobacter.  
RN NCBI\_TaxID=1061;  
[1]  
SEQUENCE FROM N.A.  
RP Burke D.H., Alberti M., Armstrong G.A., Hearst J.E.;  
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z11165; CAA77517.1; .  
DR PIR; S17805; S17805.  
KW Photosynthesis; Hypothetical protein.  
SQ SEQUENCE 55 AA; 5750 MW; 7EB55296266D48B1 CRC64;  
-----  
Query Match 12.1%; Score 36.5; DB 1; Length 55;  
Best Local Similarity 40.5%; Pred. No. 1.6e+03;  
Matches 15; Conservative 2; Mismatches 17; Indels 3; Gaps 2;  
-----  
QY 18 QPVSTRSQHTQPTPEPS--TAPSTSFLLPMGSP-PA 51  
: | | | | | | | | | | : | | | | |  
Db 9 RPSAAPARQPCGSCVTAPVAVRIGAMGASPGPA 45  
-----  
RESULT 12  
AP65\_CARMA STANDARD; PRT; 30 AA.  
AC P82964;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ANTIBACTERIAL 6.5 KDA PROTEIN (FRAGMENT).  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Decapodidae; Carcinidae;  
OX NCBI\_TaxID=6759;  
[1]  
RN SEQUENCE, AND FUNCTION.  
RP TISSUE=HEMOCYTE;  
RX MEDLINE=97008941; PubMed=8856051;  
RA Schnapp D., Kemp G.D., Smith V.J.;  
RT "Purification and characterization of a proline-rich antibacterial  
RT peptide, with sequence similarity to bactericin-7, from the haemocytes  
RT of the shore crab, Carcinus maenas".  
RL Eur. J. Biochem. 240:532-539(1996).  
CC -!- FUNCTION: STRONG ANTIMICROBIAL ACTIVITY AGAINST P. IMMOBILIS AND  
CC M. LUTEUS, LESS ACTIVE AGAINST E. COLI D22.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW IS: 6.5 KDA.  
CC -!- SIMILARITY: TO BOVINE BACTENECIN 7.  
KW Antibiotic.  
FT NON\_TER 30 30  
SQ SEQUENCE 30 AA; 3305 MW; 6E2C2205934896C4 CRC64;  
-----  
Query Match 11.9%; Score 36; DB 1; Length 30;  
Best Local Similarity 38.1%; Pred. No. 9.8e+02;  
Matches 8; Conservative 1; Mismatches 6; Indels 1; Gaps 1;  
-----  
QY 29 PTPPESTAPSTSFLLPMGSP 49  
: | | | | | | | | | | : | | | | |  
Db 4 PYPPEPRP-----PIGPRP 18  
-----  
RESULT 13

ATP8\_SQUAC STANDARD; PRT; 55 AA.  
ID Q9Z250;  
AC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).  
GN MTATP8 OR ATP8.  
OS Squalus acanthias (Spiny dogfish).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.  
OX NCBI\_TaxID=7797;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=99091711; PubMed=9873084;  
RX Rasmussen A.S., Arnason U.;  
RT "Phylogenetic studies of complete mitochondrial DNA molecules place  
RT cartilaginous fishes within the tree of bony fishes.";  
RL J. Mol. Evol. 48:118-123(1999).  
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.  
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y18134; CAA77053.1; .  
DR InterPro; IPR001421; ATP-synt\_8.  
DR Pfam; PF00895; ATP-synt\_8; 1.  
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
FT TRANSMEM 6 26 POTENTIAL.  
SQ SEQUENCE 55 AA; 6587 MW; 3FB9F843CEFA54EE CRC64;  
-----  
Query Match 11.9%; Score 36; DB 1; Length 55;  
Best Local Similarity 35.3%; Pred. No. 1.7e+03;  
Matches 12; Conservative 2; Mismatches 12; Indels 8; Gaps 1;  
-----  
QY 14 VHLPPQVSTRSQHTQPT-----PEPSTAPST 39  
: | | | | | | | | | | : | | | | |  
Db 22 VILPKKVMTHLNNPTAKSAEKPKPEPWNWPT 55  
-----  
RESULT 14  
INEL\_HUMAN STANDARD; PRT; 51 AA.  
ID O15225;  
AC 15-JUL-1999 (Rel. 38, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE INACTIVATION ESCAPE 1 PROTEIN (DXS6974E).  
GN INEL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=97386586; PubMed=9244435;  
RX Esposito T., Gianfrancesco F., Ciccodicola A., D'Esposito M.,  
RA Nagaraja R., Mazzarella R., D'Urso M., Forabosco A.;  
RT "Escape from X inactivation of two new genes associated with DXS6974E  
RT and DXS7020E.";  
RL Genomics 43:183-190(1997).  
[2]  
RN REVISIONS.





Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	60.5	20.0	52	4	Q14441	Q14441	homo sapien
2	55.5	18.4	42	4	Q16469	Q16469	homo sapien
3	52	17.2	52	12	Q92313	Q92313	human resp
4	51	16.9	52	12	Q92301	Q92301	human resp
5	51	16.9	52	12	Q92314	Q92314	human resp
6	50.5	16.7	33	6	Q02832	Q02832	gorilla gor
7	50.5	16.7	54	11	Q9WT77	Q9WT77	mus musculus
8	50	16.6	52	6	Q9GL38	Q9GL38	bos taurus
9	49.5	16.4	54	12	Q9J7C8	Q9J7C8	simian virus
10	49	16.2	46	4	Q15218	Q15218	homo sapien
11	49	16.2	51	6	Q18723	Q18723	macaca fusc
12	49	16.2	52	12	Q92302	Q92302	human resp
13	49	16.2	52	12	Q92303	Q92303	human resp
14	48	15.9	38	4	Q9HB17	Q9HB17	homo sapien
15	48	15.9	42	6	Q9GKJ3	Q9GKJ3	sus scrofa
16	47.5	15.7	27	10	Q9S8M0	Q9S8M0	solanum tub
17	47.5	15.7	57	4	Q13697	Q13697	homo sapien
18	47.5	15.7	5	5	Q9VQ06	Q9VQ06	drosoophila
19	47	15.6	31	2	Q9Z516	Q9Z516	strepomyces

AC	Q164697	01-NOV-1996	(TREMBLrel. 01, Created)
DT		01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT		01-NOV-1996	(TREMBLrel. 08, Last annotation update)
DE	GLYCOPROTEIN IB ALPHA VARIANT B (FRAGMENT).		

OX	NCBI_TaxID=11250;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=181691;
RA	Coggins W.B., Lefkowitz E.J., Sullender W.M.:
RT	"Genetic Variability among Group A and Group B Respiratory Syncytial
RL	Viruses in a Children's Hospital.";
RT	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF086872; AAC42992.1; -
DR	InterPro; IPR000925; Glycoprot.G.
DR	Pfam; PF00802; Glycoprotein_G; 1.
DR	NON_TER 1
FT	NON_TER 52
SQ	SEQUENCE 52 AA; 5613 MW; 7C3114ACA02574E6 CRC64;
Query Match	16.9%; Score 51; DB 12; Length 52;
Best Local Similarity	34.3%; Pred. NO. 73; Mismatches 4; Indels 0; Gaps 0;
Matches 12; Conservative	4; Mismatches 15; Indels 0; Gaps 0;
QY	15 HLPQVPSRSHOHTPEPSTAPSTSLFLPMGPSP 49
DB	13 HTSQKELHETTSSEGNPSQVYTTSYLQSQPSP 47
RESULT 5	
O92314	PRELIMINARY; PRT; 52 AA.
ID O92314	
AC O92314	
DT 01-NOV-1998 (TEMBLrel. 08, Created)	
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)	
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)	
DE ATTACHMENT GLYCOPROTEIN G (FRAGMENT).	
OS Human respiratory syncytial virus.	
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;	
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.	
OX NCBI_TaxID=11250;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=203721;	
RA Coggins W.B., Lefkowitz E.J., Sullender W.M.:	
RT "Genetic Variability among Group A and Group B Respiratory Syncytial	
RL Viruses in a Children's Hospital.";	
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AF086888; AAC43008.1; -	
DR InterPro; IPR000925; Glycoprot.G.	
DR Pfam; PF00802; Glycoprotein_G; 1.	
DR NON_TER 1	
FT NON_TER 52	
SQ SEQUENCE 52 AA; 5554 MW; 3B328FBC5B4E9858 CRC64;	
Query Match	16.9%; Score 51; DB 12; Length 52;
Best Local Similarity	31.8%; Pred. NO. 73;
Matches 14; Conservative	5; Mismatches 25; Indels 0; Gaps 0;
QY	6 TRSMGCAVHLPPVPSRSHOHTPEPSTAPSTSLFLPMGPSP 49
DB	4 TNSITGNLEHTSGETLHSTSSSEGNPSQAYTTSYLQSQPSP 47
RESULT 6	
O02832	PRELIMINARY; PRT; 33 AA.
ID O02832	
AC O02832;	
DT 01-JUL-1997 (TEMBLrel. 04, Created)	
DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)	
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)	
DE HUNTINGIN PROTEIN (FRAGMENT).	
GN IT15.	
GO Gorilla gorilla gorilla (Lowland gorilla).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Gorilla	

OX NCBI\_TaxID=9595;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96326790; PubMed=8766138;  
RA Pecheux C., Gail A.L., Kaplan J.C., Dode C.;  
RT "Sequence analysis of the CAG triplet repeats region in the Huntington  
RL disease gene (IT15) in several mammalian species.";  
RL Ann. Genet. 39:81-86(1996).  
DR EMBL; S83377; AAB50771.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 33 AA; 3393 MW; 44F58CE5636722A CRC64;

Query Match 16.7%; Score 50.5; DB 6; Length 33;

Best Local Similarity 38.5%; Pred. No. 54;

Matches 15; Conservative 1; Mismatches 12; Indels 11; Gaps 3;

QY 16 LPQVSTRSQHTQ-TPEPSTAPSTSELLPMGPSPPAEG 53

Db 2 LPQP----PHGQPLLPQOPQP-----PPPPPPPPG 30

RESULT 7

Q9WTY7  
ID Q9WTY7 PRELIMINARY; PRT; 54 AA.

AC Q9WTY7;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)

DE ENDOTHELIAL NITRIC OXIDE SYNTHASE (FRAGMENT).

GN NOS3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV129;

RX MEDLINE=99096466; PubMed=9878824;

RA Teichert A.M., Karantzoulis-Pegaras F., Wang Y., Mawji I.A., Bei X.,

RA Gnanapandithen K., Marsden P.A.;

RT "Characterization of the murine endothelial nitric oxide synthase

RT promoter.";

RL Biochim. Biophys. Acta 1443:352-357(1998).

DR EMBL; AF091262; AAD22613.1; -.

FT NON\_TER 54

SQ SEQUENCE 54 AA; 5193 MW; 1DA456A21958B2EA CRC64;

Query Match

Best Local Similarity 47.8%; Score 50.5; DB 11; Length 54;

Matches 11; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 29 PTEPSTAPSTSELLPMGPSPPA 51

Db 34 PAPEPQAQA-----PPSPTRPA 51

RESULT 8

Q9GL38  
ID Q9GL38 PRELIMINARY; PRT; 52 AA.

AC Q9GL38;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE CALPASTATIN (FRAGMENT).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;

OX Bovidae; Eutheria; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RA Chung H.Y., Davis M.E., Hines H.C.;  
RT "PCR-SSCP analysis of the bovine calpastatin gene domain L region.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
FR EMBL; AY008267; AAG23869.1; -.  
FT NON\_TER 1  
FT NON\_TER 52  
SQ SEQUENCE 52 AA; 5749 MW; CBD5A7449AFDDA89 CRC64;

Query Match 16.6%; Score 50; DB 6; Length 52;

Best Local Similarity 33.3%; Pred. No. 94;

Matches 11; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

QY 5 PTRMAPGAVHLPQPVSTRSQHTQTPPEPSTAP 37

Db 19 PKHSDTGTSHAPKEKAVSKSSEQPPSEKSTKP 51

RESULT 9

Q9J7C8

ID Q9J7C8 PRELIMINARY; PRT; 54 AA.

AC Q9J7C8;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE LARGE T ANTIGEN (FRAGMENT).

OS Simian virus 40 (SV40).

OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.

OC NCBI\_TaxID=10633;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MC-028863B-1;

RA Rizzo P., Carbone M.;

RT "SV40 from 1955 commercial parenteral (Salk) poliovaccine.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF180738; AAF28272.1; -.

FT NON\_TER 1

SQ SEQUENCE 54 AA; 5868 MW; 556CDAB682C1EFCD CRC64;

Query Match

Best Local Similarity 32.7%; Score 49.5; DB 12; Length 54;

Matches 16; Conservative 5; Mismatches 25; Indels 3; Gaps 2;

QY 8 SNAPGAVHLPQP-VSTRSQHTQTPPEPSTAPSTSELLPMGPSPPAEGST 55

Db 8 SQSQGFGAQPQPSQSSQSHDHNPYHICRGXTCXKKP--PTPPPPPET 54

RESULT 10

Q15218

ID Q15218 PRELIMINARY; PRT; 46 AA.

AC Q15218;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)

DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)

DE SALIVARY PROLINE-RICH PROTEIN 2 (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84298176; PubMed=6089212;

RA Azen E.A., Lyons K.M., McGonigal T., Barrett N.L., Clements L.S.,

RA Maeda N., Vanin E.F., Carlson D.M., Smithies O.;

RT "Clones from the human gene complex coding for salivary proline-rich

RT proteins";

RL Proc. Natl. Acad. Sci. U.S.A. 81:5561-5565(1984).

DR EMBL; K02578; AAA36505.1; -.

FT NON\_TER 1

FT NON\_TER 46

SQ SEQUENCE 46 AA; 4592 MW; FCEID38D8DEDC173 CRC64;

RP SEQUENCE FROM N.A.  
RC STRAIN-184473;  
RA Coggins W.B.; Lefkowitz E.J.; Sullender W.M.;  
RT "Genetic Variability among Group A and Group B  
RT Viruses in a Children's Hospital."; Respiratory Syncytial  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RR EMBL; AF086873; AAC42993.1; -;  
RD

RP SEQUENCE FROM N.A.  
RA Pirone D.M., Fukuhara S., Gutkind S.J., Burbelo P.D.,  
RT sPecSs, small binding proteins for CDC42 proteins.";  
RL J. Biol. Chem. 0:0-0(2000).  
DR ENBL: AF286592; AAG1723.1; -.  
DR InterPro: IPR001230; Prenyln. n.  
DR POSITE: P500294; PRENYLATIN; UNKNOWN\_1.  
SQ SEQUENCE 38 AA; 4245 MW; 3B416F3C5ADF4E91 CRC64;  
SQ SEQUENCE 38 AA; 4245 MW; 3B416F3C5ADF4E91 CRC64;



Query Match 15.9%; Score 48; DB 4; Length 38;  
Best Local Similarity 46.7%; Pred. No. 1.2e+02;  
Matches 14; Conservative 2; Mismatches 6; Indels 8; Gaps 2;

QY 14 VHLPPQVSTRSQHTQTPSTAPSTSFLL 43  
| | | | | : | : | |  
Db 13 VEPQPVSL-----PTPHN--PKSQLL 34

RESULT 15  
Q9GKJ3 PRELIMINARY; PRT; 42 AA.  
AC Q9GKJ3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE MYOSIN LIGHT CHAIN KINASE (FRAGMENT).  
GN MYLK.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Van Poucke M., Verle M., Tuggle C., Chardon P., Van Zeveren A.,  
RA Peelman L.J.;  
RT "Integration of porcine chromosome 13 maps."  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF222917; AAC41130.1; -  
KW Kinase.  
FT NON\_TER 1 1  
FT NON\_TER 42 42  
SQ SEQUENCE 42 AA; 4677 MW; 78BDD867E66EF64F CRC64;

Query Match 15.9%; Score 48; DB 6; Length 42;  
Best Local Similarity 36.6%; Pred. No. 1.3e+02;  
Matches 15; Conservative 2; Mismatches 14; Indels 10; Gaps 2;

QY 14 VHLPPQVSTRS----QHTQTPSTAPSTSFLLPMGSPSP 50  
| | | | | : | | | | | : |  
Db 8 VHSPPQVDFRSVLAKGKTPKTPVPEKLP-----PPKPTTP 42

Search completed: February 12, 2002, 13:03:05  
Job time: 185 sec